



# 遠隔ビデオ環境における二者間共食コミュニケーションの分析

著者	任 海因
内容記述	筑波大学修士（情報学）学位論文・平成25年11月30日授与（31047号）
発行年	2013
URL	<a href="http://hdl.handle.net/2241/00123881">http://hdl.handle.net/2241/00123881</a>

# 遠隔ビデオ環境における二者間共食コミュニケーションの分析

筑波大学  
図書館情報メディア研究科  
2013 年 11 月  
任 海因

## 目次

第1章	はじめに.....	1
1.1	研究背景.....	1
1.2	研究目的.....	1
1.3	本論文の構成.....	1
第2章	本研究の動機と意義.....	3
2.1	コミュニケーション支援.....	3
2.3	共食会話.....	3
2.3	動作の同調現象.....	4
第3章	実験.....	6
3.1	実験デザイン.....	6
3.2	実験方法.....	6
3.2.1	対面共食場面実験.....	6
3.2.2	遠隔食事実験.....	9
3.3	データの取得と処理.....	14
3.3.1	データ取得.....	14
3.3.2	データ処理.....	15
3.3.3	分析方針.....	15
3.4	分析ツール.....	17
第4章	発話についての分析.....	18
4.1	対面と遠隔発話の分析.....	18
4.2	対面と遠隔発話の頻度.....	19
4.3	発話衝突.....	20
4.4	話者交替.....	23
第5章	視線についての分析.....	26
5.1	対面と遠隔視線平均長.....	26
5.2	対面と遠隔視線頻度.....	27
第6章	食事動作についての分析.....	28
6.1	対面と遠隔食事動作平均長.....	28
6.2	対面と遠隔食事動作頻度.....	29
6.3	対面と遠隔食事動作同調.....	30
第7章	アンケート調査.....	33
第8章	検討.....	35
8.1	対面と遠隔環境に共食発話についての検討.....	35
8.2	対面と遠隔環境に共食視線についての検討.....	35
8.3	対面と遠隔環境に共食行動についての検討.....	36
8.4	参加者の主観的な評価についての検討.....	36
第9章	まとめ.....	38
謝辞	.....	39
参考文献	.....	40

# 第1章 はじめに

## 1.1 研究背景

食事には生きてくために必要な栄養を摂取するという生理的機能が備わっており、人と人を合わせることは人間同士の日常生活において不可欠である。それに加えて共に食事をする相手との親交を深めるという社会的機能も備わっている[1]。

近年、コンピュータ技術の発展に伴い人間とのコミュニケーションを目的とした人工物が多く提案されてきている。テレビ会議・電話システムを発展させた遠隔共食システムとして応用できる。日常生活においてビデオチャットなどの遠隔コミュニケーションシステムを利用する機会が増えている。日常生活では食事をしながら会話をするのはごく普通であるので、食事をしながらの遠隔会話も増えてくると考えられる。留学生と外国人労働輸出など遠く離れた家族が映像を介して一緒に食事をするシステムは現在の社会状況の課題を解決する一つの方法と期待される。その中で、家族や友人と一緒にではなく、様々な事情で一人寂しい食事をするという状況が本人の意志とは関係なく決定されてしまう「孤食」が問題になっている。「孤食」によって家族や友人とのコミュニケーションの場が失われるだけでなく、食事の偏りの原因にもなると指摘されている[2]。

そこで本研究では対面共食場面と遠隔共食場面コミュニケーション支援にとっても、どのように支援をすればよいかという本質的な問題に対して示唆を与えるものと考えられる。

## 1.2 研究目的

我々は、本研究は対面共食と遠隔共食場面のコミュニケーションを特徴を比較分析し、とくに遠隔共食と対面共食の発話、視線と食事動作に着目し、共食場面のコミュニケーションを比較分析の支援を目的とする。

この結果を受け、本研究では対面共食と遠隔共食場面のコミュニケーションを特徴を明らかにするために、対面と遠隔共食条件における参加者の発話交替、視線交錯に注目し分析を進めた。次に、対面と遠隔食事動作がどのように違いが見られるかどうか注目した。参加者の共食場面食事動作のデータから分析した。その結果、対面と遠隔共食条件に食事動作差があるとはいえず、なお、対面と遠隔環境に食事動作の同調に差があることが分かった。

## 1.3 本論文の構成

2章では関連研究を挙げながら、本研究の動機と意義について説明する。3章では、実験デザインと対面、遠隔共食場面実験について説明する。まだは、本研究で用いるデータの取得と処理について説明する。4章では発話についての分析、5章では視線についての分析分析結果を示す。6章では食事動作関連研究分析結果を、7章では参加者に注目した

アンケート調査分析結果を示す。8 章ではこれらの結果を検討する。9 章はまとめである。

## 第2章 本研究の動機と意義

本研究は、対面と遠隔環境による2者共食場面のコミュニケーションを特徴を比較分析しにはどのような違いがみられる。本章ではその動機と意義について述べる。

### 2.1 コミュニケーション支援

対面对話と遠隔映像対話の比較については、単純な会話だけでなく教示する側とされる側の共同作業といった場面におけるコミュニケーションについての比較が行われている。O' Malley らは、2者による共同作業の際のコミュニケーションについて対面環境と映像と音声を用いた遠隔環境という2つの条件間で比較している[5]。その結果、遠隔環境では対面環境よりも視線を相手に向ける時間の割合が高く、相手の発話に割り込むような発話の回数が多いことが示唆された。

### 2.2 共食会話

近年、食事中のコミュニケーションについての研究がなされている。対面環境と遠隔環境には食事の環境に心理的な距離感など。このことを明らかにするには、食事は会話、視線、食事動作に何らかの差異を及ぼすと人々に本当に思われているのかを確認する必要がある。その結果はコミュニケーション支援に役立つと考えられる。

井上らは食事の有無がコミュニケーションに与える影響について分析し、食事がある場合には参加者間の発話量が平準化するという現象を明らかにした[2]。

古川らは2者による対面共食場面と遠隔共食場面のコミュニケーション行動を比較し、さらに遠隔共食場面において互いの食事が見える場合と見えない場合を比較した[9]。その結果、食事が見えない場合には発話がより短くなり、発話衝突がより多く発生し、また、視線を相手に向ける頻度が高くなることがわかった。遠隔共食コミュニケーションにおいては食事を見せることで、より対面状況に近づくことが示唆された。

武川らは3者の共食会話を対象として、話者、次話者、非次話者という参与役割と、視線や食事行動の関係について分析している[10]。観察から、話者から次話者への視線行動は食事による影響をあまり受けないが、非次話者は食べ物を見ている頻度が高いことが分かった。また人は他人の行動と自分の行動を照らし合わせて、次に食べるべきか話すべきかを判断していること、食事では次話者の選択は発話交替の瞬間に決まるのではなく、その前からすでになされている可能性があることが示唆されたとしている。

また武川らは食事の提供形態が銘々皿の場合と大皿の場合とで、視線や食事行動が参与役割に応じて変わるかを調べている[11]。分析項目は参与役割別の視線と食事動作である。分析結果から、食事の提供形式によってコミュニケーション行動は異なり、所定のコミュニケーション効果を得るためには食事形式を適切に選択することが示唆されるとしている。本研究は、3者間の共食会話を分析対象に含む点でこれらと一致しているが、参与役割と行動の関係ではなく、食事の有無によって会話行動がどのように変わるかに注目し、共食会話の特徴を明らかにしようとしている。

さらに武川らは、共食会話では会話を優先しつつ、食事をするという構造を明らかにし

ている[12]。

徳永らは、3者の共食会話における会話構造の分析を行っている[13]。視線行動と発話行為との関わり<sup>□</sup>や、話者交替システムに基づくターン取得の特徴を抽出している<sup>□</sup>。また、話し手の発話に対する聞き手の行動に注目し、聞き手が話す、食べる行動の仕組みを分析している<sup>□</sup>。分析から、聞き手の発話、摂食のタイミングはともに、話し手の発話の区切りに適切に配置されており、聞き手は応答する、傾聴を示すなど相手とのコミュニケーションを摂食より優先することで共食会話を形成しているとしている。本研究は話し手・聞き手という役割ではなく、共食している各参加者それぞれの発話・食事行動の兼ね合いに注目している。

中山らは円滑なコミュニケーションを阻害する要因を探るために、5～6人の「飲み会」や「パーティ」を収録し、参加者の「性格テスト」「日本語能力テスト」を実施している。性格、日本語能力、グループ内の立場と発話量の関係に注目して分析し、これらいずれもが発言頻度に影響を与えることを確認している[14]。この研究は、積極的にコミュニケーションを仲介しユーザが話しやすい環境を作るメディアを設計するという目標で進められている。共食会話の分析だけでなく、その先の支援を目標としている点で、本研究はこれと一致しているが、食事の有無による会話行動を比較することによって、共食会話の特徴をあぶり出そうとしている。

## 2.3 動作の同調現象

われわれのことばと身体動作を対象とした研究は、McNeillによって[15]、個人内の認知処理過程と関連づけられるようになった。さらに近年、個人内のみならず、個人間でどのような相互作用が起こっているかが重要であることも、認知科学の分野で指摘されている[16]。こうした個人間相互作用の中でもっとも劇的なものに、身体動作のタイミングがぴったり合ってしまう現象がある。

何かの拍子に互いに同じジェスチャーをしてしまうという経験は、誰しも持っているだろう。複数の参与者間で、似た形態の身体動作が同時に起こる現象、すなわち身体動作の同期現象(Simultaneous Gestural Matching : SGM) は、限られた年齢層や文化に特有の現象というわけではない。Condon & Sander [17] は養育者の発話と乳児間で身体運動の同期が見られることを記述している。成人どうしの身体動作が同期する現象は文化が異なっても観察される[18]。また、動作の同期は人間ばかりでなく、チンパンジーにおいても確認されている[19]。

ジェスチャー研究において重要な概念に、「ジェスチャー単位」と「フェーズ」がある。身体動作には、ことばの語や文節などにあたるはっきりとした統語形式はない。日常会話で起こる自発的ジェスチャーのちょっとした腕や手の動きには、一見すると、構造らしい構造がないように見える。しかし、Kendon [20] は、ジェスチャーをいくつかの構造に分節できることを示し、その基本単位として「ジェスチャー単位」とその内部構造である「フェーズ」を定義した。「ジェスチャー単位」とは、身体の各部位がリラックスした状態から振り上げ、振り下ろし、一時停止などを経て再びリラックスした状態に戻る一連の動きを指す。ジェスチャー単位の内部はいくつかのフェーズに分かれる。腕を一振りする場合を例にとると、リラックスした位置から腕を振り上げるのが「準備preparation」、振り下ろすのが「ストロークstroke」、振り下ろした腕がいったん停止するのが「ホールドhold」、そして再びリラックスした位置に戻るのが「復帰retract」で、この一連の動作が「ジェスチャー単位」である。腕が簡単に復帰せずにあちこちをストロークによって

移動して「ジェスチャー句gesture phrase」を形成する場合もある。ジェスチャー単位の内部構造の発せられるタイミングは発語と密接な関係にあることが知られており[21]、両者の時間構造を記述することで、マルチモーダルな行為分析を行うことができる[]。



## 第3章 実験

本研究は対面と遠隔共食場面のコミュニケーションを特徴を比較分析し、食事に伴う会話については知られていないが、対面と遠隔で違いがあることが予想される。参加者は上下関係のない友人同士である。本研究は特に環境と性差に着目するものではないため、環境と性別による統制は行わなかつた。2 者間の遠隔共食場面と対面共食場面を撮影して分析を行う。

### 3.1 実験デザイン

本実験は同じ部屋での対面共食、遠隔共食、という 2 つの条件で実験を行う。この実験に半分ペアは対面条件を先に行うかについてはカウンターバランスをとる。また参加者に 2 つの条件全てに同じ相手と参加してもらう参加者内計画で実験を行う。

被験者：友人同士の大学生 20 組 40 名（男性 20 人、女性 20 人）。

### 3.2 実験方法

#### 3.2.1 対面共食場面実験

実験は筑波大学春日キャンパス講義棟 2 階の隣接した 2 つの教室に映像を通して食事ができる環境を作り（図 3）、11:00~13:00 や 17:00~19:00 の時間帯で実施する。実験中はドアとカーテンを閉め、教室の外の様子が気にならないようにする。

2 人が同じ部屋で対面する形で食事をする様子を撮影する。参加者同士の距離は友人同士では一般的とされる 120 cm とする[22]（図 1）。

正面から参加者を撮影する映像一番重要なので、視線と食事動作のビデオ分析を提供するために、ビデオカメラ (SONY HDR-CX560V 2 台) で正面からの参加者の映像を取得し、Sanyo xacti2000 1 台で横から参加者半身画面を撮影する。正面を撮る映像は解像度は 720 x 480 pix、フレームレートは 30.00fps である横から参加者半身画面を撮影する映像の解像度は 640 x 480 pix、フレームレートは 30.00fps である。撮影のフォーマットは MPG4。1 回の撮影時間は約 20 分である。

分析にはビデオ分析ツール ELAN4.3.3 を使用して、参加者ごとに会話、視線と食事動作についてラベリングする。ビデオ画面を切り離し、拡大してみる場合、映像の解像度は 640 x 480 pix 以上の設定が妥当。それから、ビデオカメラ撮る映像は解像度は 640 x 480 pix 以上の設定する。

台湾の国立中央大学「認知神経科学研究所」の研究結果を示した[23]。人間の目の視線運動形態は「凝視」と「移動」がある。眼球の「凝視」と「移動」の時間は主な指標。一般的に凝視の時間は 0.15 s ~ 0.5 s 間、平均的にの時間は 0.25 s。視線移動の平均時間は 0.035 s。実験の参加者を撮影するビデオカメラのフレームレートは 30.00fps、言い換えれば、実験の参加者の眼球移動の時間は 0.033 s 以上の目の動きを捕らえられる。視線のビデオ分析を提供する。

SONY HDR-CX560V の正面からの参加者と皿の映像を取得する。ビデオカメラ (SONY

HDR-CX560V) がら皿までに距離は150cm、参加者までに距離は160cm。撮る映像範囲は150cm～160cm、撮る映像画面のサイズは160cm x 120cm。正面を撮る映像は解像度は720 x 480 pix、毎ピクセルは画面を0.222センチメートル占めている。

Sanyo xacti2000 の撮る映像は机の上に皿と横から参加者を撮影する。ビデオカメラ (Sanyo xacti2000) がら皿までに距離は155cm、参加者までに距離は165cm。撮る映像範囲は155cm～165cm、撮る映像範囲は155cm～165cm、撮る映像のサイズは160cm x 120cm。正面を撮る映像は解像度は640 x 480 pix、毎ピクセルは画面を0.25センチメートル占めている。

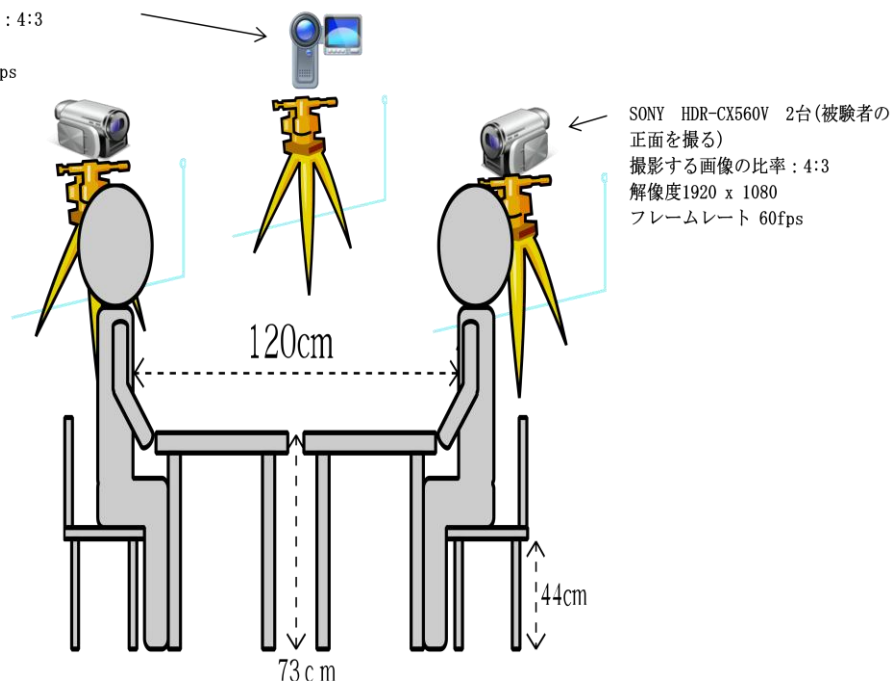
人類の眼球幅は平均2.3cm。このような条件に眼球運動のことで視線移動の方向を確認することができる。人類の口の平均幅は約8cm。実験はスプーンの長は18cm。このような条件に食物を持って、スプーンを口に入れて、口を咀嚼して様子を確認することができる。視線と食事動作のビデオ分析を提供する。

SONY HDR-CX560V の解像度はFX 1920 x 1080 pix (60.00fps) と FH 1920 x 1080 pix (60.00fps) この範囲でも適用。でもこの解像度は保存のサイズは重いから、分析するのに都合よくない。

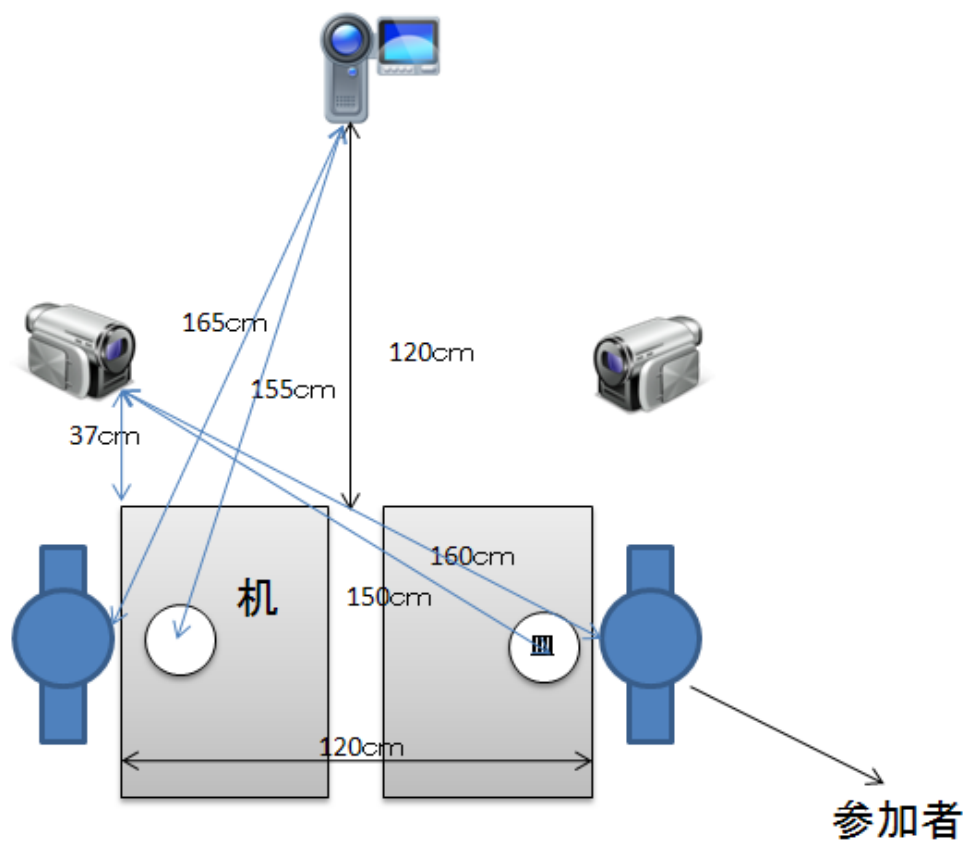
Sanyo xacti2000 の撮る映像解像度は、FHR 1920 x 1080 pix (60.00fps)、FH 1920 x 1080 pix (60.00fps)、この範囲で適用。でもこの解像度は保存のサイズは重いから、分析するのに都合よくない。

正面から参加者の上半身を撮影する映像は完全に、まだ、分析ビデオの人物のサイズは適当。練習実験を通ってくる、SONY HDR-CX560V と片側の距離は37cm。Sanyo xacti2000 と机の短い辺垂直距離は120cm。

Sanyo xacti2000 1台 (横から参加者全身画面を撮影する)  
撮影する画像の比率: 4:3  
解像度1920 x 1080  
フレームレート 60fps



(図1) 対面共食場面実験の配置図



(図 2) 対面共食場面実験の位置図



(図 3) 対面共食場面実験の様子

実験装置：

SONY HDR-CX560V 2 台(被験者の正面を撮る)

撮影する画像の比率：4:3

解像度 720 x 480 STD

フレームレート 30fps

カメラマイク オート

Sanyo xacti2000 1 台(横から参加者全身画面を撮影する)

撮影する画像の比率：4:3

解像度 640 x 480

フレームレート 30fps

カメラマイク オート

### 3.2.2 遠隔食事実験

参加者同士が互いに姿も声も確認できない離れた 2 地点に作り実験を行う (図 6)。

ウェブカメラは参加者同士の視線を合わせられるように人物像の目の高さ、また人物像と重ならないように画面の中心から左に 10cm ずらした位置に固定する。ウェブカメラで正面からの参加者の映像を取得し、ノート PC は両側映像伝送機器として、ノート PC 上でフルスクリーン表示したものをディスプレイにミラーリング表示する。ウェブカメラはズーム表示機能を持つ Microsoft LifeCam HD5000 を 2 台使用し、ノート PC は DELL LATITUDE D630 を 2 台使用する。横から参加者を撮る映像の解像度は 720 x 480 pix、フレームレートは 30fps。後参加者の斜め後ろから参加者の全身とディスプレイの画面を撮影する映像の解像度は 640 x 480 pix、フレームレートは 30fps。

1 回の撮影時間は約 20 分である。

画面の人物像に重ならず参加者同士の視線と大きく外れない位置から解像度 640×480 pix の USB カメラで参加者の正面の映像を取得し、それを PC でフルスクリーン表示した画面を相手側のディスプレイに表示した。映像の表示サイズは 827.3mm×620.5mm、解像度は 1024×768 pix、フレームレートは 30fps であった。ピクセルは画面を 0.081 センチメートル占めている。表示する人物の映像は等身大映像とし、参加者同士の距離を対面共食と同じように 120 cm とするために、画面と参加者の距離を 120 cm とした。全条件で参加者の目の高さを一致させた。映像を表示するディスプレイは Pioneer の PDP-504 CMX と NEC LCD-V462 を使用する。

分析にはビデオ分析ツール ELAN4.3.3 を使用して、参加者ごとに会話、視線と食事動作についてラベリングする。ビデオ画面を切り離し、拡大してみる場合、映像の解像度は 640 x 480 pix 以上の設定が妥当。それから、ビデオカメラ撮る映像は解像度は 640 x 480 pix 以上の設定する。

SONY HDR-CX560V の横から参加者全身画面を撮影する。撮る映像範囲は 140cm～170cm、撮る映像のサイズは 160 cm x 120 cm。正面を撮る映像は解像度は 720 x 480 pix、毎ピ

クセルは画面を 0.222 センチメートル占めている。(図 4)

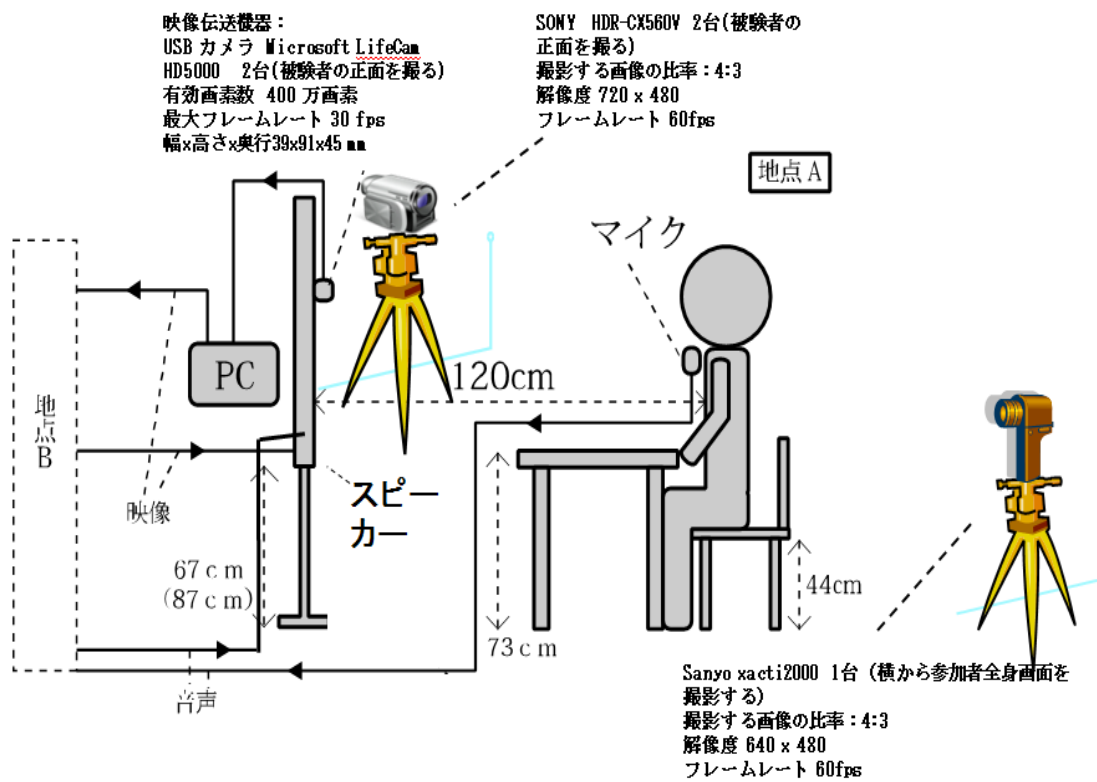
Sanyo xacti2000 斜め後ろから参加者の全身とディスプレイの画面を撮影する映像は補助分析映像とする。撮る映像範囲は 100cm~200cm、撮る映像のサイズは 160 c m x 120 c m。正面を撮る映像は解像度は 640 x 480 pix、毎ピクセルは画面を 0.25 センチメートル占めている。

SONY HDR-CX560V の解像度は FX 1920 x 1080 pix (60.00fps) と FH 1920 x 1080 pix (60.00fps) この範囲でも適用。でもこの解像度は保存のサイズは重いが、分析するのに都合よくない。

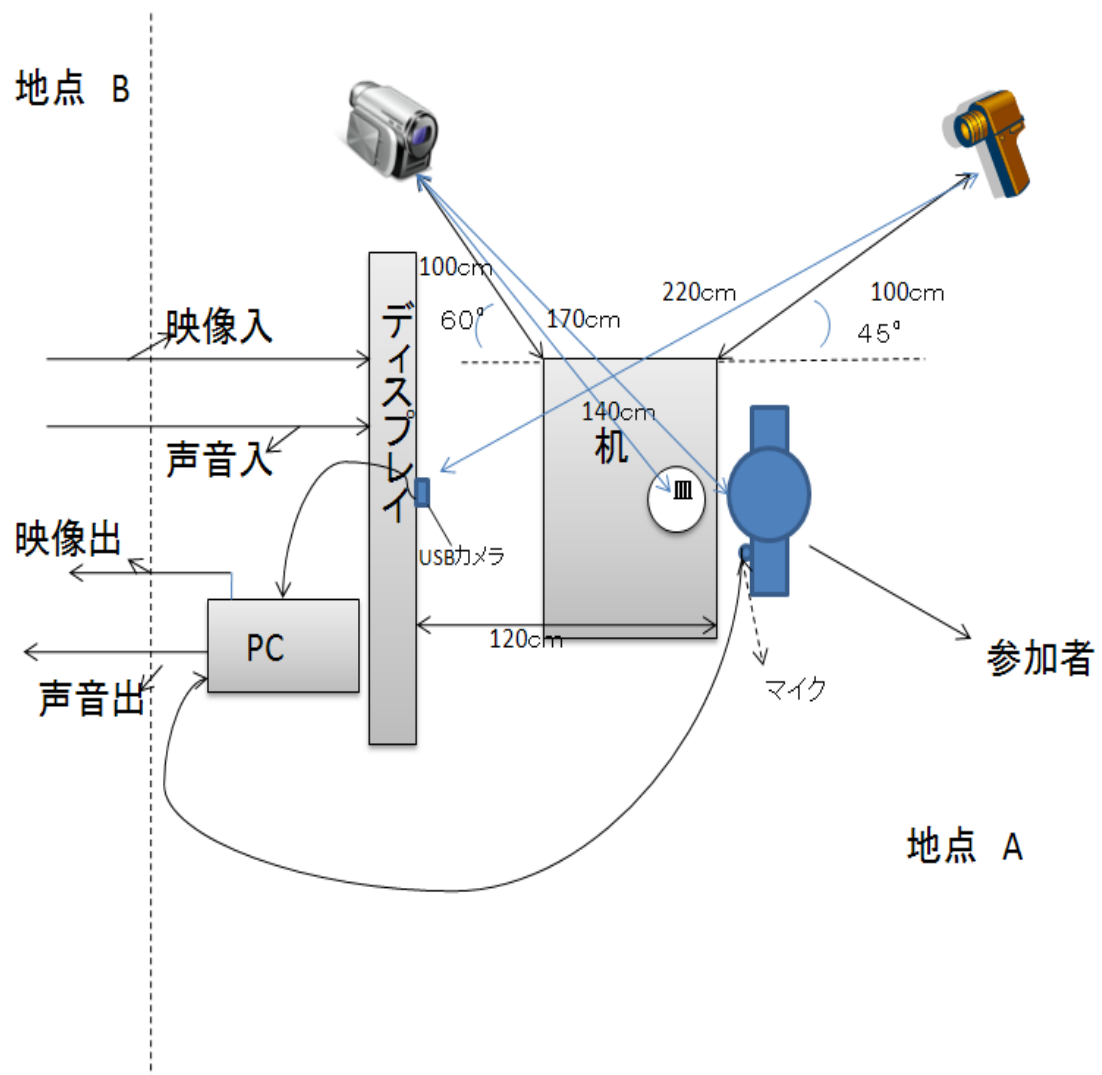
SONY HDR-CX560V の撮る映像は 140cm~170cm 範囲内に、解像度は FX 1920 x 1080 pix (60.00fps) と FH 1920 x 1080 pix (60.00fps) この範囲でも適用。Sanyo xacti2000 の撮る映像は 100 c m ~ 220 c m 範囲内に、FHR 1920 x 1080 (60.00fps)、FH 1920 x 1080 pix (60.00fps)、この範囲でも適用。でもこの解像度は保存のサイズは重いが、分析するのに都合よくない。横から参加者の上半身を撮影する映像は完全に、まだ、分析ビデオの人物のサイズは適当。練習実験を通ってくる、SONY HDR-CX560V と片側の距離は 100cm、机の短い辺と夾角は 60°。斜め後ろから参加者の全身とディスプレイの画面を撮影する映像は Sanyo xacti2000 と机の短い辺は 100cm、夾角は 45°。

コミュニケーションに関する先行研究では対人距離を 120cm として実験を行なっているため、本実験では人物像を等身大に表示した画面と参加者との距離を 120cm とする。(図 4)

音声はマイクとスピーカーを用いて参加者が問題なく会話できる音量と音質で伝送する。マイクは SONY ECM-719 を 2 台使用し、音声については、参加者同士が支障なく会話できる音量と音質のマイクとスピーカーを使用する。テレビのスピーカーを使用するので、正面音声の出所を提供することができる。音量調節も数量化の実現のために助けを提供する。



(図 4) 遠隔共食場面実験の配置図



(図 5) 遠隔共食場面実験の位置図



(図 6) 遠隔共食場面実験の様子

実験装置：

テレビ：Pioneer PDP-504 CMX と NEC LCD-V462

Pioneer PDP-504 CMX 液晶：50 型

映像入力 1 チャンネル

音声入力 1 チャンネル

映像のサイズは 827.3mm×620.5mm、

解像度は 1024×768、(MAX1440×900)

フレームレートは 30.00fps

NEC LCD-V462 液晶： 50 型

映像入力 1 チャンネル

音声入力 1 チャンネル

映像のサイズは 827.3mm×620.5mm、

解像度は 1024×768、(MAX1440×900)

フレームレートは 30.00fps

映像伝送機器：

Microsoft LifeCam HD5000 2 台(被験者の正面を撮る)

有効画素数 400 万画素

最大フレームレート 30fps

幅 x 高さ x 奥行 39x91x45 mm

映像伝送ソフトウェア Debut Video Capture Software で設定はフォーカスは自動して、  
ズームは 2 階する。



映像伝送ノート PC : は DELL LATITUDE D630 を 2 台

CPU Core2Duo T7100 1.8GHz

メモリ 2048MB

HD 80GB

OS Microsoft Windows 7

USB ポート 4 ポート USB2.0 対応

映像伝送ソフトウェア : Debut Video Capture Software (有料ダウンロード)

撮影機器 :

SONY HDR-CX560V 2 台 (横から参加者を撮影する) ,

撮影する画像の比率 : 4:3

解像度 720 x 480 STD

フレームレート 30fps

カメラマイク オート

Sanyo xacti2000 2 台, (参加者の斜め後ろから参加者の全身とディスプレイの画面を撮影する)

撮影する画像の比率 : 4:3

解像度 640 x 480

フレームレート 30fps

音声伝送 :

スピーカー NEC SP-RM2 2 台

定格出力 15W

インピーダンス 8 ohm

マイク :

被験者のえりにピンマイクをつける 直流バイアス式コンデンサー型マイク SONY ECM-719 2 台

周波数: 100Hz-15kHz

電池型番 LR44

体積 : 6.2 x 3.8 x 1.1

重量 : 1.3 パウンド

### 3.3 データの取得と処理

#### 3.3.1 データ取得

自然な会話のデータが取得できるように話題や食事を終える時間は指定せず、普段通りに振舞ってもらうように指示をする。食事は 2 人で「頂きます」と言って始めてもらう。食事の種類による会話への影響を除くため、全ての参加者の食事をカレーライスとお茶、使う食器をスプーンに統一する。

対面条件では個人撮影用 1 台ずつと全景撮影用 1 台、合計 3 台のビデオカメラで撮影す

る。遠隔条件では個人撮影用2台ずつ、合計4台のビデオカメラで撮影し、USBカメラで取得する参加者の正面からの映像についても録画する。

参加者への指示と食事の用意をしてから撮影を開始し、自由なタイミングで食事を始めてもらう。2人とも食べ終わったことを確認してから撮影を終了する。参加者の食事中、実験者は部屋の外で待機した。1回の撮影時間は約12分～26分であった。

### 3.3.2 データ処理

本研究では食事をしているときのコミュニケーションを対象としているため、撮影した映像はMPG、対面共食場面に被験者の正面を撮る映像は解像度は720 x 480 pix、横から参加者全身画面を撮影する映像の解像度は640 x 480 pix。遠隔共食場面に被験者の正面を撮る映像の解像度は640 x 480 pix、横から参加者を撮影する映像の解像度は640 x 480 pix、参加者の斜め後ろから参加者の全身とディスプレイの画面を撮影する映像の解像度は640 x 480、フレームレートは30.00fps、1回の撮影時間は約20分である。撮影した映像は動画編集ソフトを使用して食事開始の数秒前から2人とも食べ終わった数秒後までをトリミングし、後述する会話分析ソフトで使用するためAVI形式に変換する。解像度は640×480、フレームレートは30.00fpsである。

### 3.3.3 分析方針

本研究では言語的側面として「会話」に着目する。食事の場面を考慮している先行研究[2]について、すべての条件で自然な会話のデータが取得できるように、特に話題は指定せず、普段通りに振舞ってもらうように指示をする。会話の場面を考慮している先行研究[13]どちらの条件でも自然な会話のデータが取得できるように話題は「安楽死の是非」、「同性同士の結婚認否」、「卒業後の進路」、「卒業旅行に行きたい場所」、「好きなテレビ番組は何か」、「好きな映画について雑談」「最近に何をおいしい食べ物を食べた」。参加者同士は自然な会話をたに、最初の設定され話題に2～3話題を自由に選択することができる。また、中国人の会話は、日本人の会話と同じように分析して、比較することが目的であるため、非言語的側面として「視線」に、さらに対面共食場面と遠隔共食場面の食事があるときの特徴的な行動として「食事動作」に着目し分析を行う。2者間の対面共食場面と遠隔共食場面のコミュニケーション行動を比較し、行動の転換を分析し、その原因を明らかにする。分析にはビデオ分析ツールELANを使用して、参加者ごとに会話、視線と食事動作についてラベリングする。ELAN4.4.3[24]はビデオファイルに注釈を加える為のプログラム。映像、音、音波画面を時系列ともに表示、再生することができ、さまざまなフォーマットにも対応しているため、分析などに使用することに優れている。ビデオ画面を切り離し、拡大してみる場合、映像の解像度は640 x 480 pix以上の設定が妥当。フォーマットはAVIやWAV。

また取得した映像について観察したところ、数人の参加者について食事を開始した直後などに部屋の外を気にしたり、教室内をきょろきょろと見回したりするなど、実験環境に戸惑うような素振りが見受けられる。そのため、そのような行動が落ち着いてある程度の時間が経ったと判断した食事開始から60秒後までの区間は分析対象外とする。分析時間は各ペアのそれぞれの条件についてまず300秒間予定とする。

**発話行動の分析方法：**

発話行動のタグ付けの際の基本単位として、「0.1 秒以上の無音区間によって区切られた、単一話者の連続する音声区間」を発話単位とする間休止単位に基づいて、音声区間を「発話」タグ付する。

また、「発話」の中には会話の内容に関わる通常の発話の他に、話者に対する了承や話を続けて欲しいというシグナルとして機能する「うん」や「そう」というように単独でのあいづち、発話の意思はあるが思考中であるというときなどに発する「あー」や「えーと」といった間を埋める言葉であるフィラーというように単独でのあいづち、心理的同調性を示しあいづちの代わりとしても機能する笑いといったものがある。参加者の発話行動をそれぞれ性質の異なる 4 種類の発話と、沈黙を合わせた 5 種類に分けてタグ付けする[2]。

発話行動についての分析。1 分間に何回それぞれの発話と沈黙が起こっているかという発話頻度と 1 回の発話と沈黙の平均長の分析。

本研究では参加者にタスクを与えているわけではないが、対面と遠隔条件に差があるため、発話数に何らかの傾向が見られるのではないかと考える。

まず、各条件の発話時間と沈黙時間について比較し、なんらかの傾向が調査する。発話衝突についての分析。取得した映像を観察したところ、2 人が同時に発話を開始してしまう回数に条件間で差があるように感じたため、同時発話の回数について分析を行う。ここでは 0.2 秒以内の差でほぼ同時に開始した通常の発話を「発話衝突」としている。発話衝突が起こった直後には参加者の一方もしくは両方が自分の発話を中断するという行動が見られる。

#### 視線方向の分析方法：

視線は感情や相手への態度の表現、情報収集、会話の流れを調整する機能を備えている。共同作業時には相手の顔だけでなく作業環境や手元を見たり、食事中には自分の食事や他人の食事を見たりしている。

本研究では食事中の視線について、自分の食事・相手・相手の食事・その他の 4 種類に分類しタグ付けする。なお、遠隔共食場面と対面共食場面に視線平均長と頻度。

#### 各条件に話者の食事動作起こる分析方法：

対面と遠隔食事中食事行動タグ付する。本研究はに Home position (Ho)：「食事に関与していない状態」、Hold tableware before eating (Htb)：「食物を取っている状態」、Hold foods (Hf)：手に持った「料理を保持している状態」、Eating (E)：「料理を口に入れている状態」、Hold tableware after eating (Hta)：「口から手を下ろしている状態」、Hold drink (Hd)：「コップを手に持っている状態」、Drinking (D)：「飲み物を飲んでいる状態」をと分類しそれぞれタグ付した。

#### アンケート分析方法[2]：

参加者の主観的な評価について取得するために、各条件の実験後、評定尺度法によるアンケートに回答してもらった。評定をそれぞれ数値に置き換え、全く当てはまらない=1、当てはまらない=2、あまり当てはまらない=3、どちらとも言えない=4、やや当てはまる=5、当てはまる=6、よく当てはまる=7 という 7 段階で評価してもらった。質問項目については、22 項目設けた。各条件においてコミュニケーション行動が違うことが予想されるが、例えば発話の頻度などに違いがあった場合、参加者がそれを意識できたかどうか、意図的にそうしたのかについて 12 項目で評価してもらった。実際に行われたコミュニケーションについても評価してもらう必要があるため、各条件での会話の楽しさやしやすさについて、それによって相互理解できていたかについて評価してもらった[4]。それに加えて、本実験では食事について扱っているため、食事の「楽しさ」や「しやすさ」についても 4 項目で評価してもらった。さらに、実験環境が適切であったかについて、映像と音

声という 2 項目について評価してもらった。

順序効果による偏りをなくするため、質問項目の順序は毎回ランダムにした。アンケートの最後には、「実験に参加して感じたこと」について自由記述で回答してもらった。

### 3.4 分析ツール

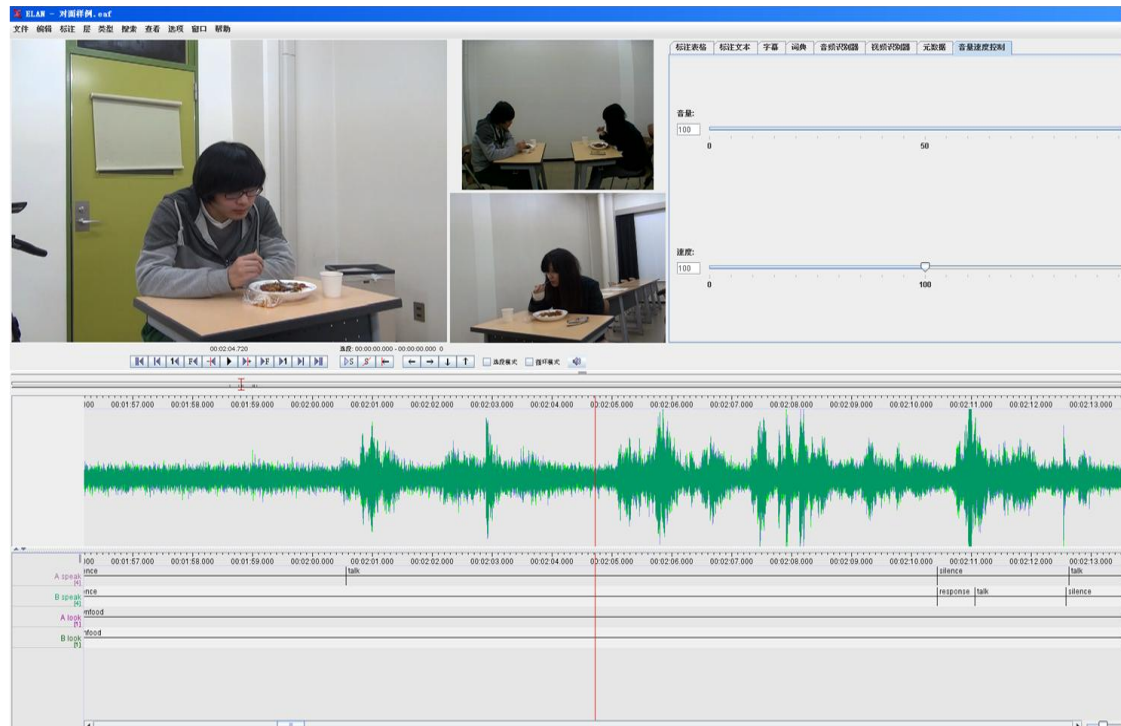


図 7 ELAN4. 3. 3 の作業画面

撮影した映像は、ビデオ分析ツール ELAN4. 3. 3[24]を使用して、参加者ごとに発話とジェスチャ、食事行動についてラベリングした。ELAN4. 3. 3 は、同時に複数映像を見ながらタイムラインに対して複数種類のラベリングを出力できる。図 7 は ELAN4. 3. 3 の作業画面である。

対面と遠隔食事条件それぞれ、各参加者につき 300 秒分、合計 12,000 秒をラベリングした。

## 第4章 発話についての分析

発話のタグ付けの際の基本単位として、「0.1 秒以上の無音区間によって区切られた、単一話者の連続する音声区間」を発話単位とする間休止単位[25]に基づいて、音声区間を「発話」、無音区間を「沈黙」としてタグ付した。

また、「発話」の中には会話の内容に関わる通常の発話の他に、話者に対する了承や話を続けて欲しいというシグナルとして機能する「うん」や「そう」というように単独でのあいづちや、発話の意思はあるが思考中であるというときなどに発する「あー」や「えーと」といった間を埋める言葉であるフィラー、心理的同調性を示しあいづちの代わりとしても機能する笑いといったものがある[26]。参加者の発話をそれぞれ性質の異なる4種類の発話と、沈黙を合わせた5種類に分けてタグ付けした。「発話」の中には会話の内容に関わる「通常の発話」の他に、話者に対する了承や話を続けて欲しいというシグナルとして機能する「あいづち」、会話の間を埋めたり発話の意思はあるが思考中であるというシグナルとして機能したりする「フィラー」、心理的同調性を示したり、あいづちの代わりになったりする「笑い」といったものがある[2]。例えばフィラーや笑いが相手の発話と重なったとしても、会話の流れにさほど影響がないなどそれぞれ性質が異なるため、本研究では発話の種類を分けて分析を行った。「うん」や「そう」といった単独で発話された「あいづち」を<response>、「あー」や「えーと」といった間を埋める言葉や言いよどみといった「フィラー」を<filler>、発声の伴う「笑い」を<laugh>、その他の「通常の発話」を<talk>というように、4種類に分類してタグ付した。「沈黙」は<silence>とタグ付した。

### 4.1 対面と遠隔発話の分析

表 1 対面と遠隔環境に発話分類時間

環境	データ量	通常の発話	あいづち	フィラー	笑い	沈黙	総発話時間
対面	40 (人)	3451.48 (秒)	179.070 (秒)	104.655 (秒)	137.362 (秒)	8032.527 (秒)	12000 .000 (秒)
遠隔	40 (人)	3245.610 (秒)	171.153 (秒)	81.928 (秒)	152.315 (秒)	8521.854 (秒)	12000 .000 (秒)

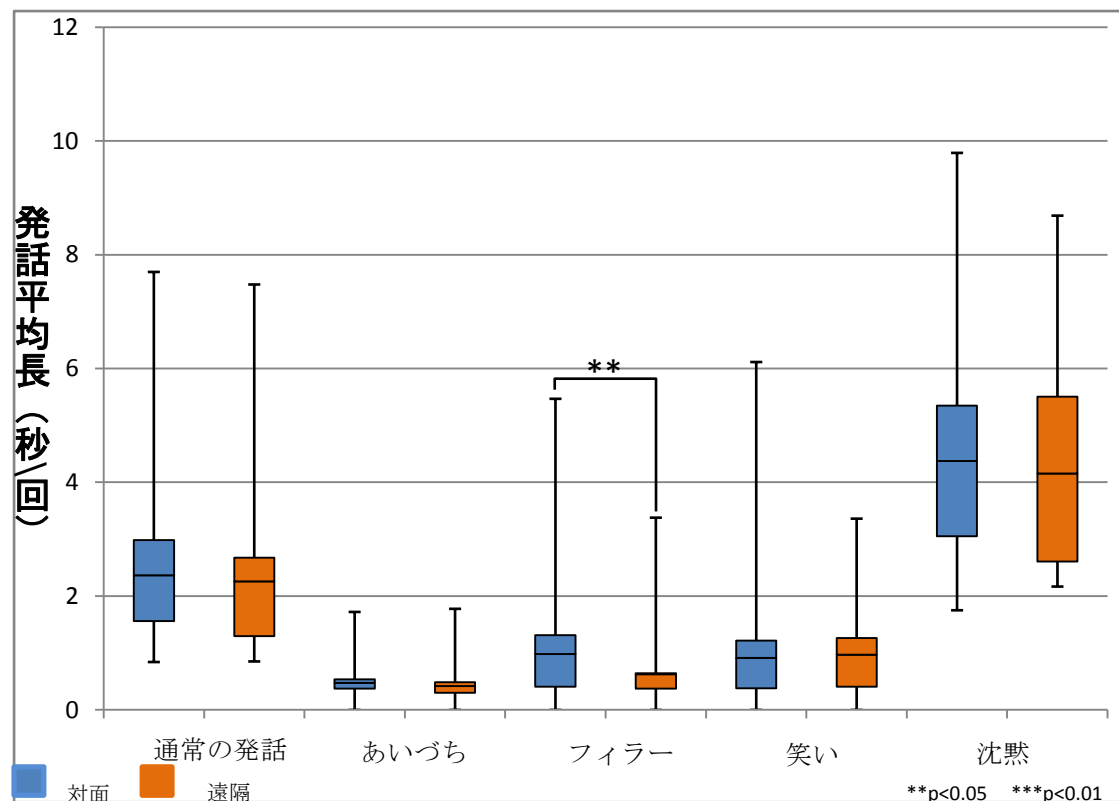


図 8 対面と遠隔発話平均長の図

米村らは、遠隔環境において低スキルのユーザがオペレータとのやり取りを通してタスクをこなすという場面では、映像がある場合の方が音声のみの場合に比べて発話数が少なくなると述べている[27]。本研究では参加者にタスクを与えているわけではないが、遠隔環境に差があるため、発話時間に何らかの傾向が見られるのではないかと考えた。

図 8 は対面環境に発話平均長を示す。正規分布しているとは言えない対応のあるデータであったため、各発話についてデータは正規分布をなすかどうか分からない場合に Wilcoxon の符号付順位和検定を行った。フィラーは ( $N=20$ ,  $z=2.10$ ,  $p<0.05$ ) 条件の間で有意差が認められた。対面環境では遠隔環境よりフィラー時間が長いことが分かった。

## 4.2 対面と遠隔発話の頻度

表 2 対面と遠隔環境に発話分類回数

環境	データ量	通常の発話	あいづち	フィラー	笑い	沈黙	総発話数
対面	40 (人)	1720 (回)	421 (回)	94 (回)	147 (回)	2220 (回)	4602 (回)
遠隔	40 (人)	1687 (回)	429 (回)	144 (回)	176 (回)	2376 (回)	4812 (回)

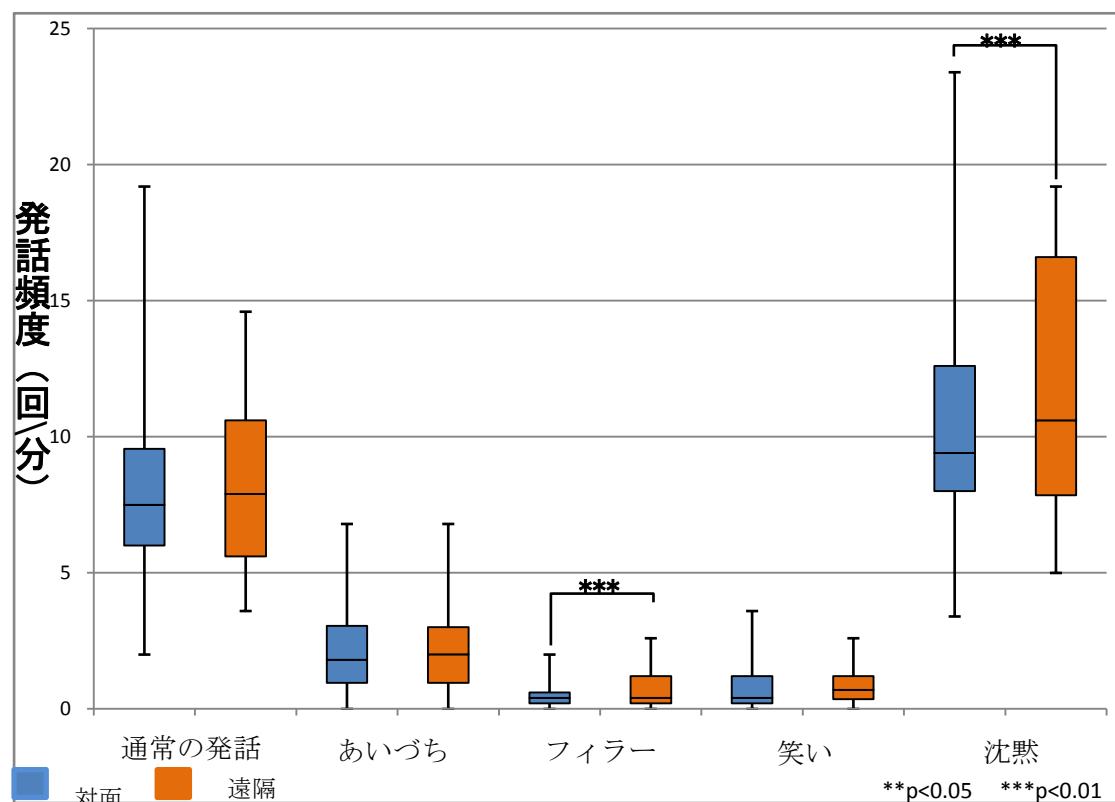


図9 対面と遠隔発話頻度の図

図9対面と遠隔環境に各発話の頻度を示したものである。各発話についてデータは正規分布をなすかどうか分からない場合にWilcoxonの符号付順位和検定を行った。フィラー (N=20,  $z=4.04$ ,  $p<0.01$ )、沈黙 (N=20,  $z=5.09$ ,  $p<0.01$ ) 条件の間で有意差が認められた。対面環境では遠隔環境よりフィラーと沈黙頻度が低いことが分かった。

### 4.3 発話衝突

遠隔共食では、対面共食より発話衝突が多く起こることがわかっている[28]。古川らは話者の発話や視線による話者交替規則を提案し、発話衝突はその規則から外れる現象としている[9]。また小磯らは、発話衝突を非円滑な話者の移行としている[29]。取得した映像を観察したところ、2人が同時に発話を開始してしまう回数に条件間で差があるように感じたため、同時に発話が開始された回数について分析を行った。ここでは2人が0.2秒以内の差[29]でほぼ同時に開始した通常発話を「発話衝突」としている。発話衝突が起こった直後には参加者の一方もしくは両方が自分の発話を中断するという行動が見られた。

表3 対面と遠隔環境発話衝突が総会話役割

環境	データ量	総会話数	発話衝突数	総会話数に役割
対面	20 (ペア)	1726 (回)	17 (回)	5.00%
遠隔	20 (ペア)	1549 (回)	36 (回)	5.00%

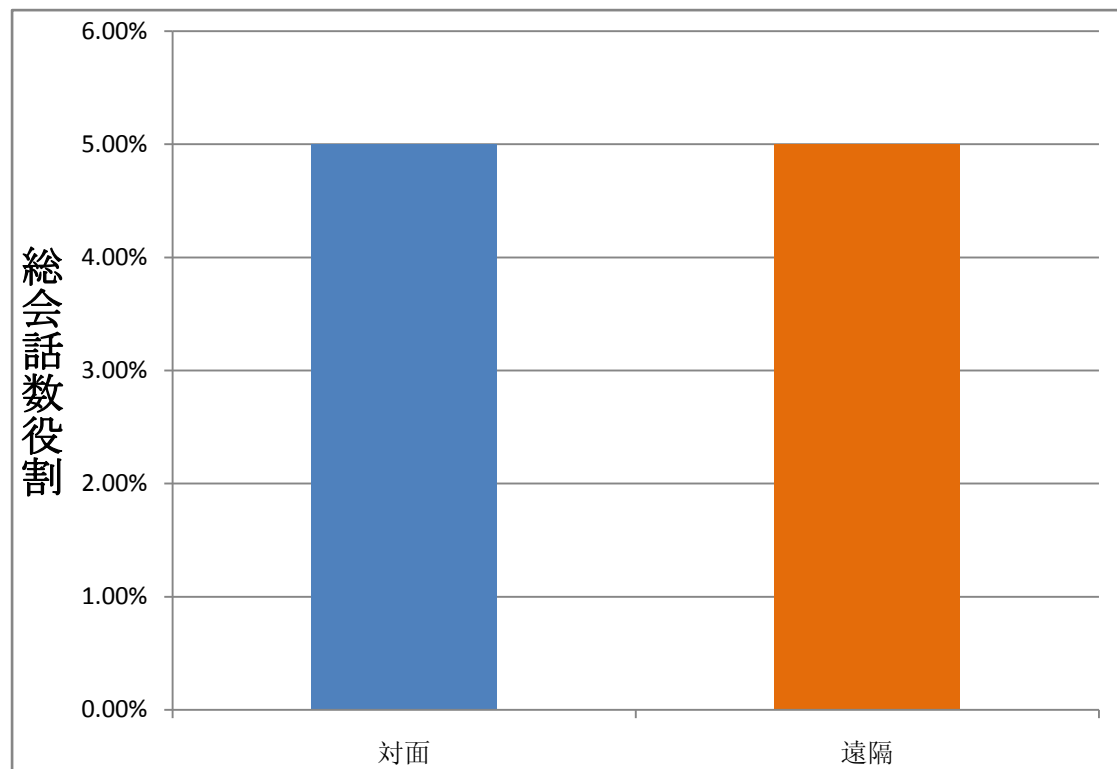


図 10 対面遠隔環境発話衝突が総会話数役割図

発話衝突が起こった回数が総会話数に役割を条件別に表したのが図12である。対面と遠隔環境に発話衝突が起こった回数が総会話数に役割が多ほとんど同じ。

表 4 対面と遠隔環境発話衝突が頻度

環境	データ量	時間	発話衝突数	発話衝突頻度
対面	20 (ペア)	100 (分)	17 (回)	0.17
遠隔	20 (ペア)	100 (分)	36 (回)	0.36



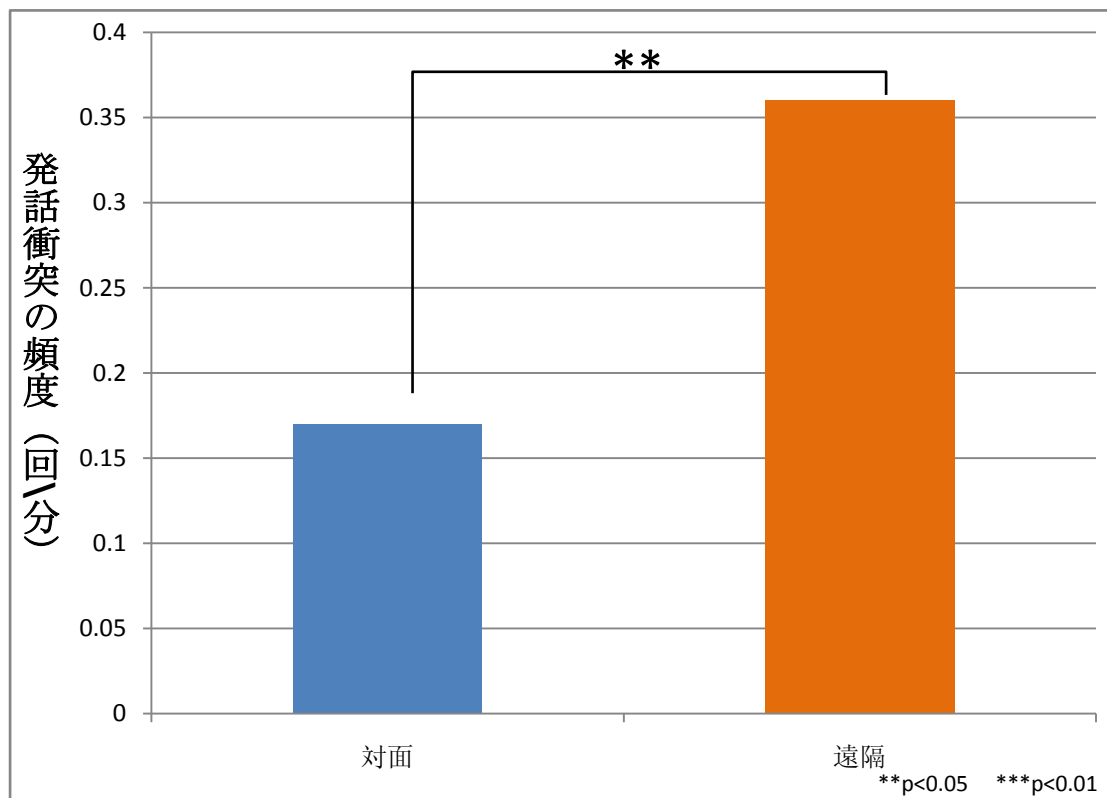


図 11 対面と遠隔環境発話衝突が頻度

発話衝突が起こった回数が総会話数に役割を条件別に表したのが図11である。各発話衝突データは正規分布をなすかどうかわからない場合にWilcoxonの符号付順位和検定を行った。対面と遠隔条件の間に有意差が認められた ( $N=10$ ,  $z=2.24$ ,  $p<0.05$ )。対面環境条件より遠隔環境に発話衝突頻度が多いことがあることが分かった。

表 5 対面と遠隔環境発話衝突が頻度

環境	データ量	総会話時間	発話衝突時間	総会話時間に役割
対面	20 (ペア)	3329.63 (秒)	62.19 (秒)	1.87%
遠隔	20 (ペア)	3056.95 (秒)	90.70 (秒)	3.28%

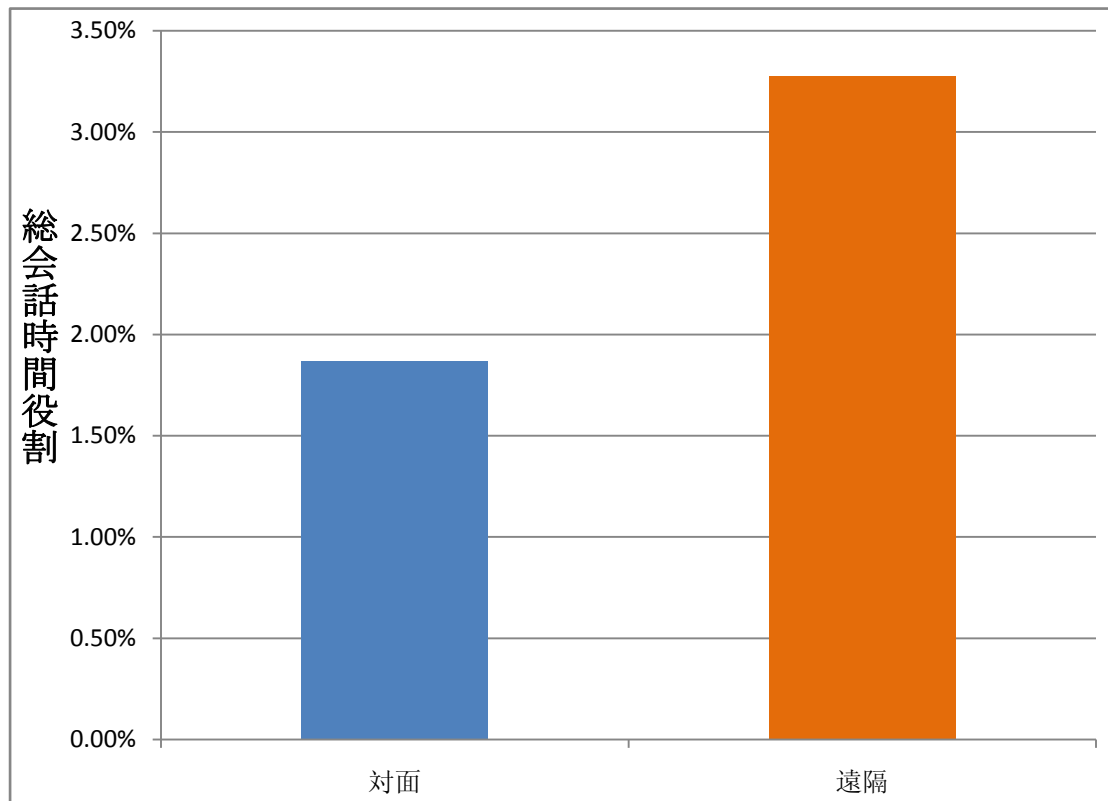


図 12 対面と隔環境発話衝突時時間の役割

対面と遠隔環境に発話衝突が起こった時間が総会話時間に役割を条件別に表したのが図12である。正規分布しているとは言えない対応のあるデータであったためWilcoxonの符号付順位和検定を行ったところ、対面環境と遠隔環境に発話衝突が起こった時間が総会話時間に役割が有意差が間で認められなかった。

#### 4.4 話者交替

榎本らは話者交替に関する定義がある[30]。各ターンの生じた時間的關係に基き、話者交替の型を以下に定義する。話者が交替するものは3種類型がある。非重複型：無音区間を挟んで話者が交替。準重複型：先行発話の終了と同時に（ $-+100\text{ms}$ 以内）に話者交替。重複型：重複を伴い話者が交替（先行話者か先に発話を終了）。話者が交替しないものは2の型がある。重複内包型：重複を生じるが先行話者が発話を継続（重複話者が先に発話を終了）相づち型：聞き手の相づちをまたいで先行話者が発話を継続。

ここでは、一方が通常の発話を開始した後に、最初に現れたもう一方の通常の発話を、話者が交替した発話とみなしカウントした。あいづちやフィラーのみを返したり、笑いによる発声の場合は話者交替とみなさなかった[31]。また、フィラーや笑いについても、はっきりと何かを話しているわけではないため、話者交替としてカウントしなかった。

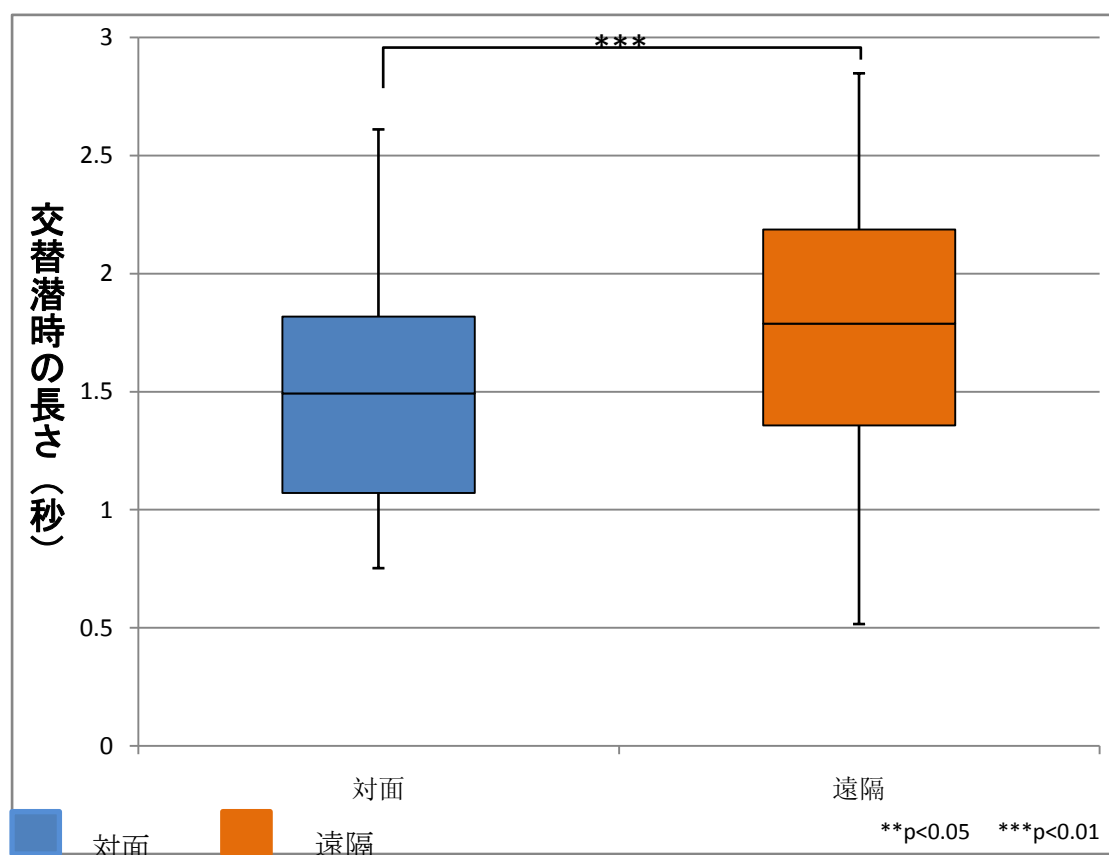


図 13 対面と遠隔環境交替潜時の長さ

図13 対面環境と遠隔環境に話者交替潜時の長さを表したものである。正規分布していない対応のあるデータであったため、Wilcoxonの符号付き順位和検定を行ったところ、対面環境に有意差が認められた ( $N=10$ ,  $z=3.90$ ,  $p<0.05$ )。従って、対面環境により遠隔環境に話者交替の間隔が長いことが分かった。

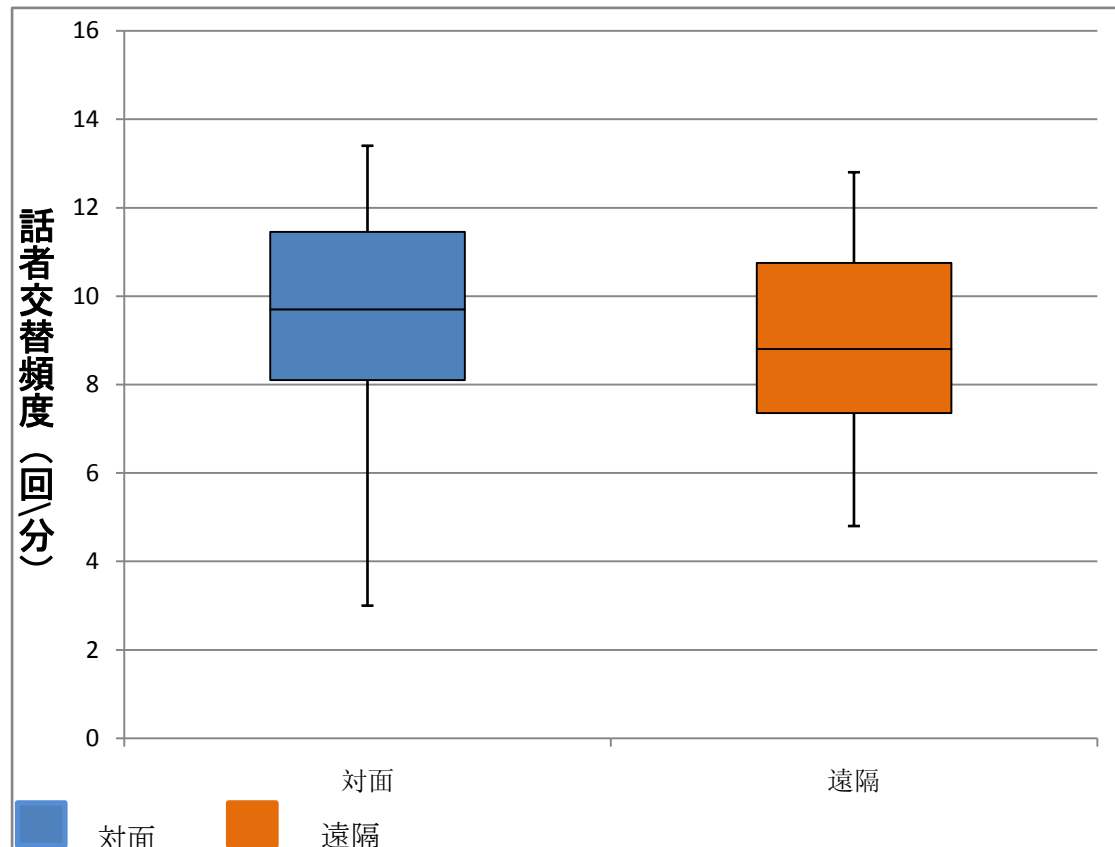


図 13 対面と遠隔環境交替潜時の長さ

¥

小宮によれば、「話し手は、聞き手の巧みな相づちにより、話を続ける意欲を得る」とされている[32]。相づちは会話を円滑に進める上で必要不可欠な要素と言ってもよいであろう。楊晶は電話会話における[33]、相づちの頻度及び話者交替に差異のあることを指摘している。

図13は対面環境と遠隔環境に話者交替の頻度を表したものである。正規分布していない対応のあるデータであったため、Wilcoxonの符号付き順位和検定を行ったところ、両方環境に条件間で話者交替の頻度に有意差は認められなかった。

## 第5章 視線についての分析

視線は非言語的コミュニケーション行動として感情や相手への態度の表現や情報収集、会話の流れを調整する機能を備えている[34]。加えて、食事場面では食事行動を行うために自分の食事を見たり、他人の食事の進行状況を確認したりする役割を果たすと考えられる[35]。

本研究では食事中の視線方向について、「相手」を見ている時間、「相手の食事」を見ている時間、「自分の食事」を見ている時間、「その他の方向」を見ている時間の4種類に分類し、それぞれ<partner>、<partner's food>、<own food>、<other>とタグ付した。

### 5.1 対面と遠隔視線平均長

表 6 対面と遠隔環境に視線分類時間

環境	データ量	自分の食事	相手	相手の食事	その他	総視線時間
対面環境	40 (人)	8367.39 (秒)	2192.12 (秒)	825.53 (秒)	267.76 (秒)	12000.00 (秒)
遠隔環境	40 (人)	7655.87 (秒)	2812.48 (秒)	856.03 (秒)	62.14 (秒)	12000.00 (秒)

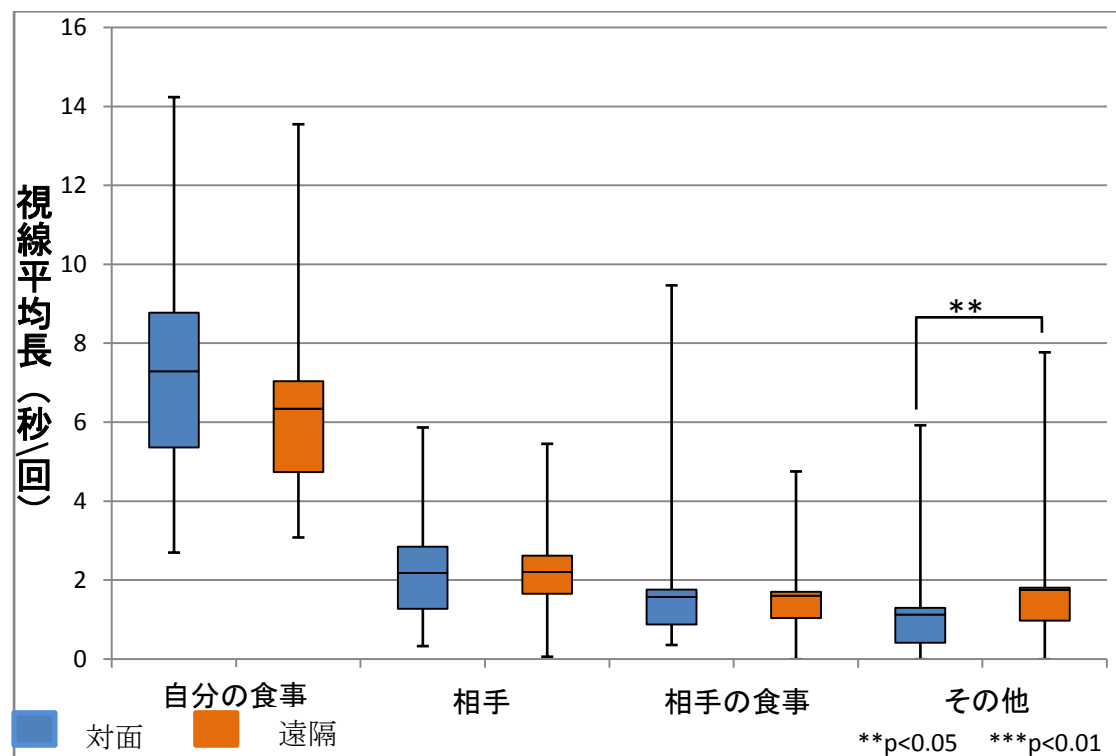


図 14 対面と遠隔環境視線平均長の図

図14は対面と遠隔環境に視線の平均長を示したものである。正規分布していない対応の

あるデータであったため、Wilcoxonの符号付き順位和検定を行ったところ、その他(N=20,  $z=2.17$ ,  $p<0.05$ )の間で有意差が認められた。対面環境では遠隔環境よりその他見る時間が短いことが分かった。

## 5.2 対面と遠隔視線頻度

表 7 対面と遠隔環境に視線分類回数

環境	データ量	自分の食事	相手	相手の食事	その他	総視線回数
対面環境	40 (人)	1219 (回)	1066 (回)	592 (回)	367 (回)	3244 (回)
遠隔環境	40 (人)	1339 (回)	1299 (回)	546 (回)	431 (回)	3456 (回)

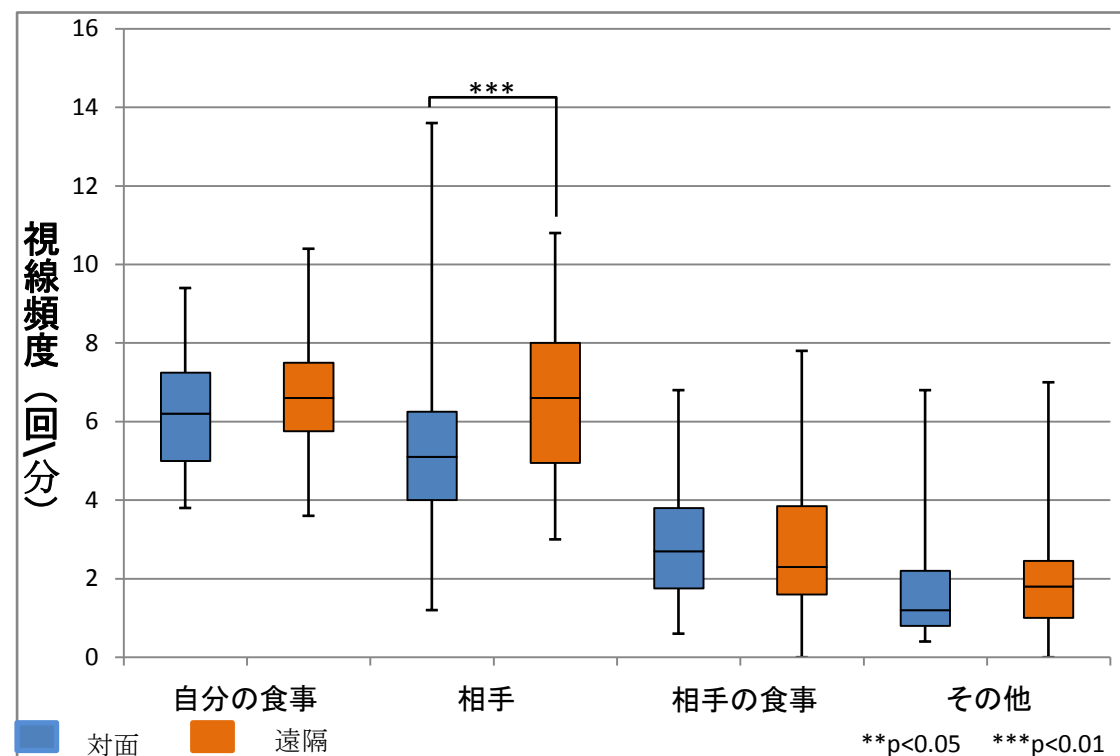


図 15 対面と遠隔環境視線の頻度の図

図 15 は対面と遠隔環境に視線の頻度を示したものである。二方条件に視線頻度を比較し、正規分布していない対応のあるデータであったため、Wilcoxon の符号付き順位和検定を行ったところ、相手(N=20,  $z=2.55$ ,  $p<0.01$ )、で有意差が認められた。対面環境では遠隔環境より相手見る頻度が低いことが分かった。

## 第6章 食事動作についての分析

大武らは、ホームポジションから準備、ストローク、ストロークの前後にホールド、復帰という動作を経てホームポジションに戻るというジェスチャフェーズをもとに、食事行動を分類している[36]。

本研究はにHome position (Ho) : 「食事に関与していない状態」、Hold tableware before eating (Htb) : 「食物を取っている状態」、Hold foods (Hf) : 手に持った「料理を保持している状態」、Eating (E) : 「料理を口に入れている状態」、Hold tableware after eating (Hta) : 「口から手を下ろしている状態」、Hold drink (Hd) : 「コップを手に持っている状態」、Drinking (D) : 「飲み物を飲んでいる状態」をと分類しそれぞれタグ付した。

### 6.1 対面と遠隔食事動作平均長

表8 対面と遠隔環境に食事動作分類時間

環境	データ量	Ho	Htb	Hta	Hf	E	Hd	D	総視食事動作時間
対面	40 (人)	2357.4 8 (秒)	3534.5 4 (秒)	588.3 5 (秒)	4155.3 2 (秒)	939.4 2 (秒)	312.1 3 (秒)	143.9 6 (秒)	12000.0 0 (秒)
遠隔	40 (人)	2455.4 6 (秒)	3497.5 1 (秒)	558.7 1 (秒)	3947.0 6 (秒)	711.8 2 (秒)	429.3 3 (秒)	160.2 2 (秒)	12000.0 0 (秒)

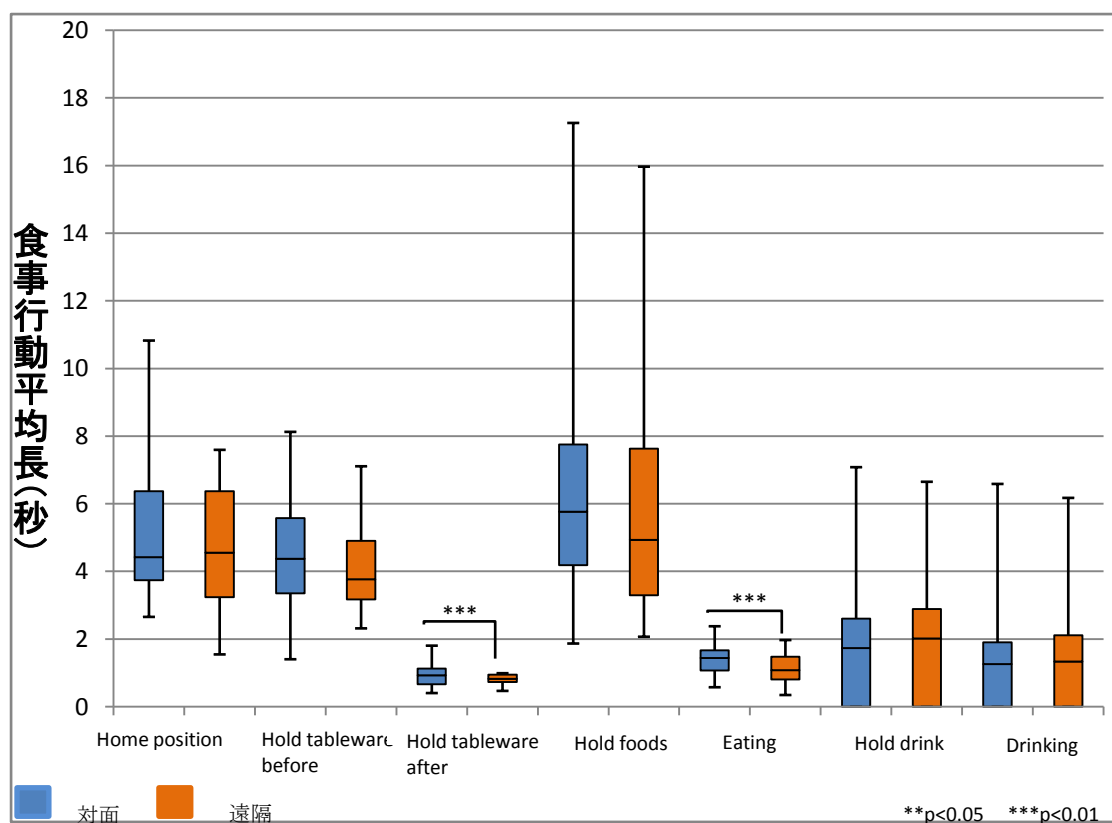


図 16 対面と遠隔環境食事動作平均長の図

図 16 は対面と遠隔環境に食事動作平均長を示したものである。二方環境に食事動作平均長を比較し、正規分布していない対応のあるデータであったため、Wilcoxon の符号付き順位和検定を行ったところ、口から手を下ろしている状態（ $N=20$ ,  $z=2.60$ ,  $p<0.01$ ）と料理を口に入れている行動（ $N=20$ ,  $z=1.50$ ,  $p<0.01$ ）平均長が有意差が認められた。対面環境では遠隔環境より口から手を下ろしている状態と料理を口に入れている行動時間が長いことが分かった。

## 6.2 対面と遠隔食事動作頻度

表 9 対面と遠隔環境に食事動作分類回数

環境	データ量	Ho	Htb	Hta	Hf	E	Hd	D	総視食事動作回数
対面	40 (人)	448 (回)	788 (回)	640 (回)	748 (回)	683 (回)	124 (回)	63 (回)	3494 (回)
遠隔	40 (人)	486 (回)	880 (回)	694 (回)	753 (回)	659 (回)	147 (回)	85 (回)	3704 (回)



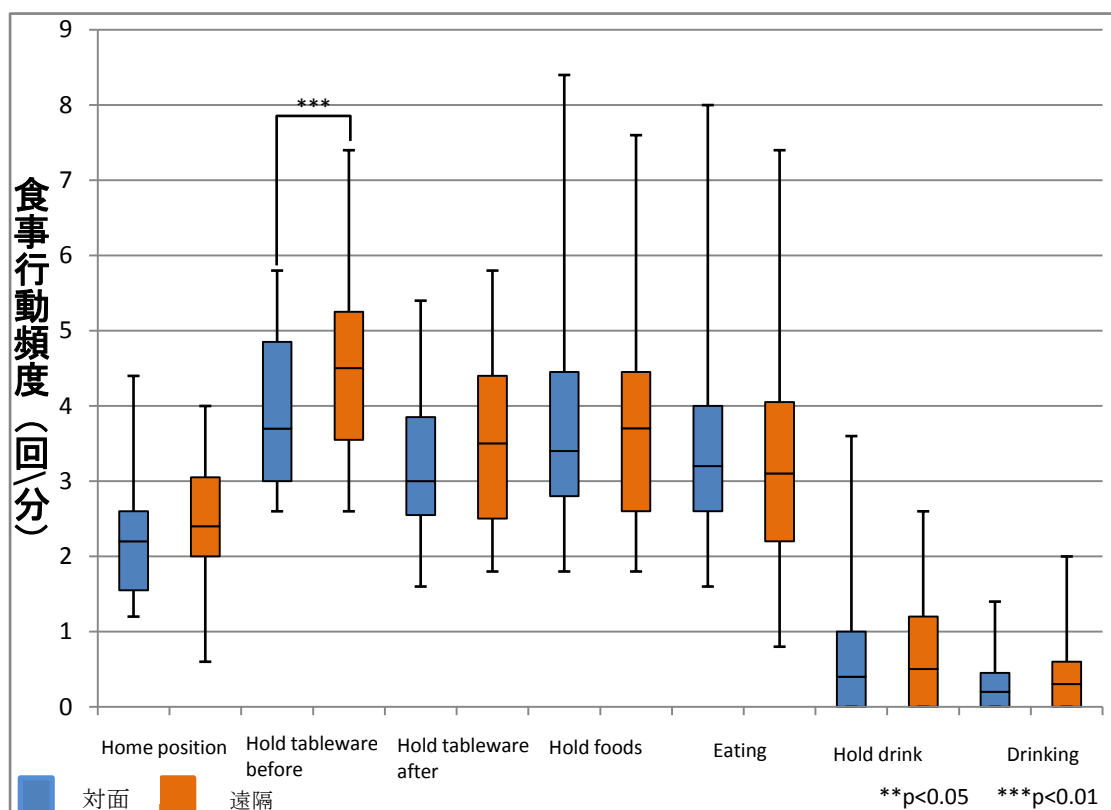


図 17 対面と遠隔環境食事動作頻度の図

図16は対面と遠隔環境に食事動作頻度を示したものである。二方環境に食事動作平均長を比較し、正規分布していない対応のあるデータであったため、Wilcoxonの符号付き順位と検定を行ったところ、食物を取っている状態（ $N=20$ ,  $z=2.58$ ,  $p<0.01$ ）頻度が有意差が認められた。対面環境では遠隔環境より口から食物を取っている状態行動時頻度が低いことが分かった。

### 6.3 対面と遠隔食事動作同調

身体動作の同期は、あらかじめ社会的・文化的に決まった定型のジェスチャー(エンブレム)でのみ起こるのではない。形態がより自由で個人差の多いはずの自発的ジェスチャー[37]でも、条件が整えば、形態とタイミングが見事に一致することがある。このような同期は、単なる偶然の産物とは考えにくい。

そして、もし偶然でないとしたら、参加者は食事のコースにさまざまな可能性を持ちうる自発的ジェスチャーの中から、限られた形態やタイミングを選ぶ必要がある。いつ、誰と、どのような内容でジェスチャーを同期させればよいのか、その可能性を参加者たちはどのようにして絞り込むのだろうか。

各ターンの生じた時間的關係に基き、参加者の同種類食事行動をしている状態。ここでは、一方が食事行動をしている状態にもう一方の伴い同じ食事行動した食事行動時間と回数をカウントする。

表 10 対面と遠隔環境に食事動作同調分類時間

環境	データ量	Ho	Htb	Hta	Hf	E	Hd	D	総視食事動作
----	------	----	-----	-----	----	---	----	---	--------

									時間
対面	20 (ペア)	357.78 (秒)	587.30 (秒)	15.34 (秒)	680.09 (秒)	48.20 (秒)	12.62 (秒)	2.69 (秒)	1704.01 (秒)
遠隔	20 (ペア)	422.56 (秒)	557.37 (秒)	16.25 (秒)	604.79 (秒)	23.06 (秒)	8.49 (秒)	0.72 (秒)	1633.24 (秒)

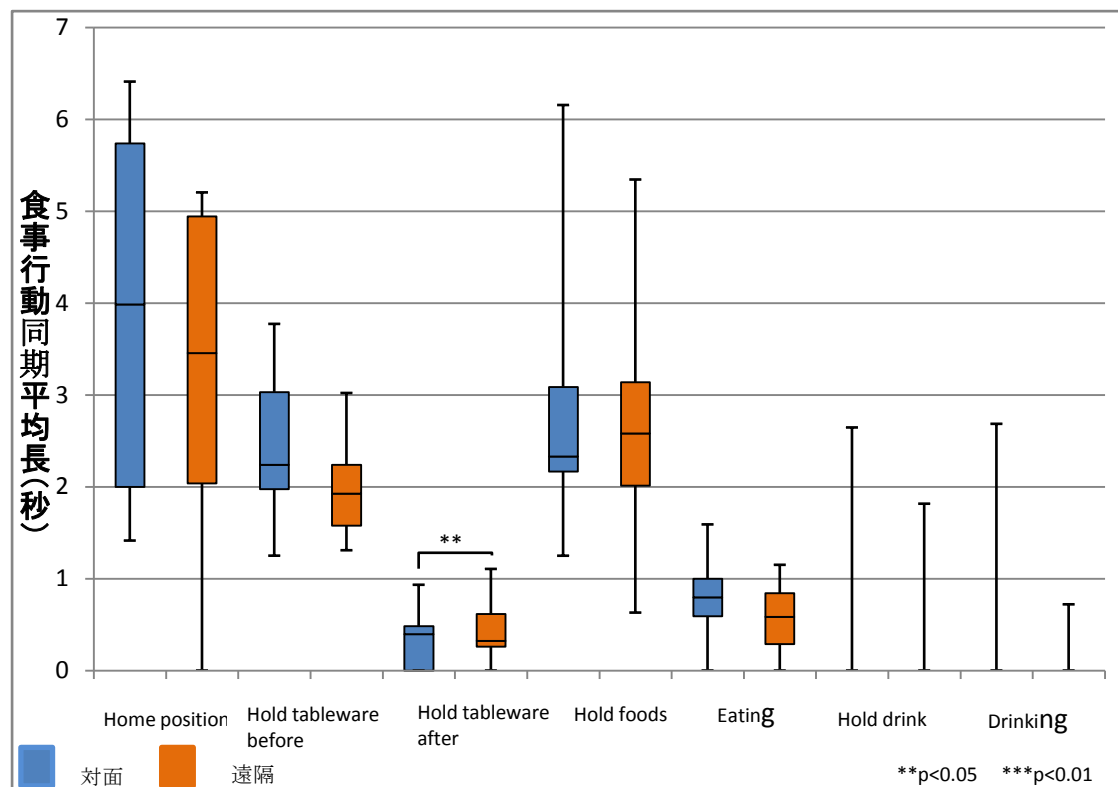


図 18 対面と遠隔環境に食事動作同調平均長の図

図18は対面と遠隔環境に食事動作平均長を示したものである。二方環境に食事動作平均長を比較し、正規分布していない対応のあるデータであったため、Wilcoxonの符号付き順位検定を行ったところ、口から手を下ろしている状態（N=20,  $z = 2.50$ ,  $p < 0.05$ ）平均長が有意差が認められた。対面環境では遠隔環境より口から手を下ろしている状態行動時間が短いことが分かった。

表 11 対面と遠隔環境に食事動作同調分類回数

環境	データ量	Ho	Htb	Hta	Hf	E	Hd	D	総視食事動作回数
対面	20 (ペア)	92 (回)	238 (回)	28 (回)	243 (回)	59 (回)	6 (回)	1 (回)	667 (回)
遠隔	40 (人)	106 (回)	264 (回)	34 (回)	233 (回)	36 (回)	9 (回)	1 (回)	683 (回)

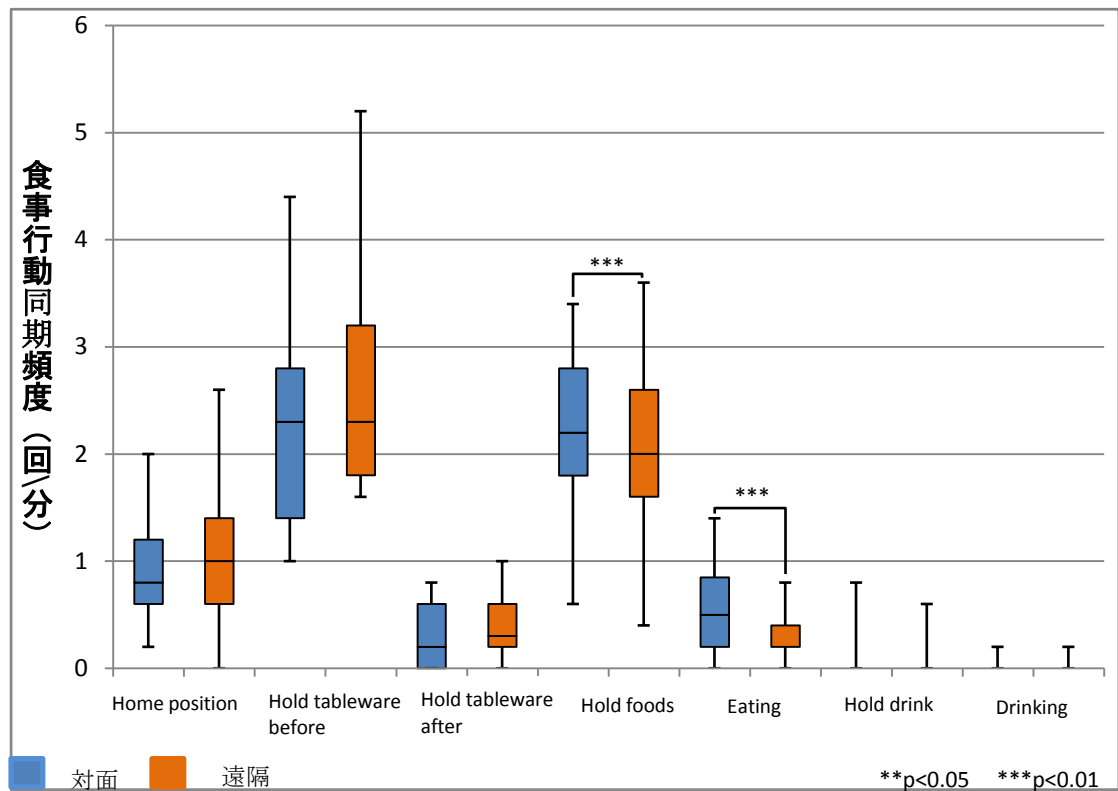


図 19 対面と遠隔環境に食事動作同調頻度の図

図19は対面と遠隔環境に食事動作頻度を示したものである。二方環境に食事動作平均長を比較し、正規分布していない対応のあるデータであったため、Wilcoxonの符号付き順位和検定を行ったところ、手に持った「料理を保持している状態」(  $N=20$ ,  $z=3.90$ ,  $p<0.01$  ) と料理を口に入れている状態 (  $N=20$ ,  $z=3.70$ ,  $p<0.01$  ) が有意差が認められた。対面環境では遠隔環境より料理を保持している状態と料理を口に入れている状態行動時頻度が高いことが分かった。

## 第7章 アンケート調査

参加者の主観的な評価について取得するために、各条件の実験後、評定尺度法によるアンケートに回答してもらった。評定をそれぞれ数値に置き換え、全く当てはまらない＝1、当てはまらない＝2、あまり当てはまらない＝3、どちらとも言えない＝4、やや当てはまる＝5、当てはまる＝6、よく当てはまる＝7 という7段階で評価してもらった。質問項目については、22項目設けた。各条件においてコミュニケーション行動が違うことが予想されるが、例えば発話の頻度などに違いがあった場合、参加者がそれを意識できたかどうか、意図的にそうしたのかについて12項目で評価してもらった。実際に行われたコミュニケーションについても評価してもらう必要があるため、各条件での会話の楽しさやしやすさについて、それによって相互理解できていたかについて評価してもらった[38]。それに加えて、本実験では食事について扱っているため、食事の「楽しさ」や「しやすさ」についても4項目で評価してもらった。さらに、実験環境が適切であったかについて、映像と音声という2項目について評価してもらった。

順序効果による偏りをなくすため、質問項目の順序は毎回ランダムにした。アンケートの最後には、「実験に参加して感じたこと」について自由記述で回答してもらった。

表 12 各条件における評定の平均値

	項目	対面	遠隔	検定の結果
1	食事を楽しむことができた	5.8	5.4	n. s
2	食事をしやすかった	5.9	5.6	n. s
3	相手の食事の様子が分かりやすかった	5.9	5.6	n. s
4	相手の食事の進み具合が気になった	4.1	4.3	n. s
5	会話を楽しむことができた	6.3	5.8	***
6	会話をしやすいと感じた	6.1	5.0	***
7	相手の言っていることがよく分かった	6.5	6.0	***
8	自分が言っていることが相手によく伝わったと思う	6.2	5.7	***
9	頻繁に発言していたと思う	4.7	5.0	n. s
10	発言が短かったと思う	4.0	4.2	n. s
11	沈黙が短かったと思う	5.3	4.5	***
12	話者が頻繁に替わったと思う	4.1	4.3	n. s
13	頻繁に発言するようにした	3.7	4.2	*
14	発言を短くしようとした	3.0	3.1	n. s
15	長い時間沈黙しないようにした	4.3	4.5	n. s
16	交互に話そうとした	5.2	5.3	n. s
17	話し始めるタイミングがよく重なったと思う	3.3	3.1	n. s
18	自分の手元をよく見ていたと思う	4.2	3.7	n. s
19	相手の顔をよく見ていたと思う	5.3	4.9	**
20	相手の手元をよく見ていたと思う	3.8	3.6	n. s

21	映像がはっきりしていた	5.5	5.0	n. s
22	音声をはっきり聞き取ることができた	5.8	4.7	***
*:p<0.1, **:p<0.05; ***p<0.01				

表 13 自由記述による回答

条件	コメント
対面	圧力がない、自然な感じ 距離感は存在しない 通りの食事だった 画面を見るよりお互いの音を理解しやすい 食事場所は教室、空の感じ、通常食事の場所が違い 相手の表情が注意する いつも通りの会話と食事だった 相手が目の前にいるのが自然でリラックスできた 画面越しより話しやすい 話題が自然に出てきた 一番話やすかった 相手の雰囲気を感じ取れて話やすかった 食事をしていると、身振り手振りをするのが難しかった
遠隔	距離感がある、非現実、特に話題を探したいと思った 画面で食事が可笑しいな感じ 通りの食事が違い、 マイクで、会話が不自然な感じ 距離感の食事が奇妙な感じ 相手の雰囲気を感じ取れて難しい 非現実感じ、特にお互いに話をしたい 画面越しに会話するのは少し恥ずかしかった いつもなら気にしないことがいつも以上に気になった 見られること、見ていることを意識した。 目があっているように感じて話やすかった 食事中はあまり相手のことを見ていない気がするが、画面越しだと相手を見るようになった 画面越しだと距離感が分かりにくかった 手元が見えると安心できた

それぞれの項目に対する条件別の評価の平均値を示す。各対面と遠隔環境のデータで、2群による比較であったため、正規分布しているとは言えない対応のあるデータであったためWilcoxonの符号付順位和検定を行ったところ、各環境に7項目について有意差が認められた。または、1項目について有意差が傾向があることが分かった。また表2は、自由記述による回答を条件別に示したものである。

## 第8章 検討

### 8.1 対面と遠隔環境に共食発話についての検討

図9より、対面と遠隔環境に発話の頻度についての分析結果から、対面環境よりフィラーの頻度が高いことが明らかになった。沈黙について、対面環境より沈黙の頻度が高いことが明らかになった。会話分析では、フィラーは発話権の保持[39]、場つなぎとしての対話潤滑話[40]、後続発話に対する知識の検索[41]の機能があるとされている。佐々木は沈黙とフィラーの関係を分析し、会話中の沈黙時にフィラーが表出されることで「積極性」の印象が高まる傾向があるという結果が得られる[42]。沈黙中に特に意味を含まない発話であるフィラーでさえそれを発することで、会話に参加しようとする姿勢を示すことができ、話そうとしている態度を場に知られる効果があるのではない。このことにより、フィラーは発話権取得機能に寄与している可能性がある。対面環境よりフィラーの頻度が高いこと図8に対面環境にフィラーが短いことが分かったということが、対面環境より遠隔環境に参加者はもっと多い発話権を取得することにつながっていると考えられる。

発話衝突についての分析の結果。図10より対面遠隔環境発話衝突が総会話数役割について、対面と遠隔環境に発話衝突が起こった回数が総会話数に役割が多ほとんど同じ。図11より対面と遠隔環境発話衝突が頻度について、対面環境条件より遠隔環境に発話衝突頻度が多いことがあることが分かった。賈は小集団討論場面において、日本語における話者交替の実態を反映できる話者交替システムを提示した。参加者は10分間に同時発話の回数が10回、割合が1.7%で事を示している[43]。表中12の17「話し始めるタイミングがよく重なったと思う」に対する評価より、評定の平均値は条件間で同じ値を示していることある。従って、対面環境と遠隔環境には、非円滑な話者の移行である発話衝突の生起を抑制し、より円滑な話者の移行を可能にするという効果が考えられる。

話者交替についての分析結果。小宮によれば、「話し手は、聞き手の巧みな相づちにより、話を続ける意欲を得る」とされている[32]。相づちは会話を円滑に進める上で必要不可欠な要素と言ってもよいであろう。楊晶は電話会話における、相づちの頻度及び話者交替に差異のあることを指摘している[33]。表中13の12「話者が頻繁に替わったと思う」に対する評価より、評定の平均値は条件間で同じ値を示していることある。これは、両方環境では、話し終えた時にすぐに自分が話し始めることが出来ない状況があることによる影響が考えられる。

### 8.2 対面と遠隔環境に共食視線についての検討

図15より相手を見る頻度が、遠隔環境では対面環境よりも高いことが明らかになった。これは、O'Malleyは遠隔環境では対面環境よりも視線を相手に向けるようになるという結果[44]と同様であった。また、表6対面と遠隔環境に視線分類時間について、共食時には相手を見る時間より自分の食事を見る時間の方が長いことが分かる。自分の食事を見ている総量時間割合にすると対面環境には約69.73%、遠隔環境には約63.80%映像の観察

からも、参加者が食事行動を行うために自分の食事を見ている様子が多く見られた。Argyle は対面環境において、視線を相手に向ける割合が 60%である事を示している[45]。コミュニケーションにおける文化差[46]についても考慮する必要があるが、共食場面では対面対話より相手を見る時間が少ないと考えられる。しかし、表 12 の 18「自分の手元をよく見ていたと思う」についての評定がどの条件でも「どちらともいえない」と「あまり当てはまらない」の間ほどにあることから、どの条件でも参加者は特に自分の食事をよく見ていることをあまり意識していない。反して、表 12 の 19「相手の顔をよく見ていたと思う」についても評定が対面環境には「やや当てはまる」と「当てはまる」の間ほどにある、遠隔環境には「どちらとも言えない」と「やや当てはまる」の間ほどにある。両方環境に有意差傾向があり。そして、遠隔環境では対面環境より参加者は相手を見ている様子が多く見られた。

### 8.3 対面と遠隔環境に共食行動についての検討

武川は食事動作を分析について、食事動作はジェスチャの分析手法[47]を簡易化し、箸や食器を持っていない「ホーム (Home position)」状態、箸、スプーン、カップなどの食器持っているがまだ食べ物を持っていない「食器把持 (Hold tableware)」状態、食べ物を持っている「食物把持 (Hold food)」 「摂取 (eating)」状態を分けて提示した。[10]。話者や聞き手にかかわらず食器を把持している状態、そのとき自分の食器と相手を見ている頻度も高い。この結果に観察している食器を把持している時自分の食器と相手を見ている頻度も高いことが可能である。本研究について、図 17 より対面と遠隔環境食事動作頻度結果から、遠隔環境には対面環境より「食物を取っている状態」の頻度が高いことが明らかになった。図 15 より、相手の見ている視線の頻度について対面環境より相手を見る同じ結果がある。

身体動作の同期は、あらかじめ社会的・文化的に決まった定型のジェスチャー(エンブレム)でのみ起こるのではない。形態がより自由で個人差の多いはずの自発的ジェスチャー(McNeill, 1992)でも、条件を整えば、形態とタイミングが見事に一致することがある。このような同期は、単なる偶然の産物とは考えにくい[44]。また、Furuyamaは、折り紙の折り方を教える場面を観察して、学習する側が教示者のストロークを追い越す事例を記述しているが、この事例では、折り紙の形状の持つ対称性や繰り返し構造などが、学習者による投射可能性を高めている可能性がある[48]。

そして、もし偶然でないとしたら、参加者は食事のコースにさまざまな可能性を持ちうる自発的ジェスチャーの中から、限られた形態やタイミングを選ぶ必要がある。いつ、誰と、どのような内容でジェスチャーを同期させればよいのか、その可能性を参加者たちはどのようにして絞り込むのだろうか。例えば、映像の観察からも、食事場面では食事行動を行うために自分の食事を見たり、相手の食事の進行状況を確認する。今後、食事動作を持っている形態と視線の関係を考慮に入れる必要があるだろう。

### 8.4 参加者の主観的な評価についての検討

アンケート調査表中の表 12 の 5「会話を楽しむことができた」と 6「会話をしやすいと感じた」に対する評価は、遠隔の 2 条件より対面条件のほうが高い。また、7「相手の言っていることがよく分かった」と 8「自分が言っていることが相手によく伝わったと思う」に対する評

価より、対面環境では遠隔環境より相手の言っていることがよく分かったと感じていることが分かる。さらに表 2 を見ると、「話題が自然に出てきた」、「相手の雰囲気を感じ取れて話やすかった」というように、対面条件ではポジティブな回答が多い。従って、対面条件でのコミュニケーションが一番好ましく感じていたと考えられる。

表 12 の 3 「相手の食事の様子が分かりやすかった」より、遠隔環境に相手の食事の様子が距離感と感じたことが分かった。また表 13 より、「距離感がある、非現実、特に話題を探したいと思った」という記述があったことから、遠隔環境と一緒に食事をしていると感じることが出来ず、食事がしにくくなると考えられる。従って、食事についても対面条件のものが一番好印象であったと考えられる。

また、表 12 の 22 「音声をはっきり聞き取ることができた」に対する評価より、本実験の遠隔環境の映像や音声の状態に問題がなかったといえる。



## 第9章 まとめ

本研究は対面共食と遠隔共食場面のコミュニケーションを特徴を比較分析し、とくに遠隔共食と対面共食の発話、視線と食事動作に着目し、対面と遠隔環境に共食場面のコミュニケーションを比較分析の支援を目的とし、遠隔共食と対面共食のコミュニケーションについて基本的な指標である発話、視線、食事動作について比較した。また、実験参加者にはアンケートに回答してもらった。

発話と視線について分析した結果、遠隔共食では対面共食より相手を見る頻度が高くなることが分かった。さらに、対面環境では遠隔環境よりフィラーが短く、発話衝頻度が低いことが分かった。また、対面環境により遠隔環境に話者交替の間隔が長いことが分かった。また参加者は対面共食においての会話や食事が好ましく感じたことが分かった。

食事動作について分析した結果、遠隔環境では対面環境より食物を取っている状態頻度が高いことが分かった。対面と遠隔環境に食事動作平均長について、遠隔環境より対面環境に口から手を下ろしている状態と料理を口に入れている行動時間が長いことが分かった。食事動作同調について分析した結果、遠隔環境では対面環境より料理を保持している状態と料理を口に入れている状態頻度が低いことが分かった。また、対面環境により遠隔環境に口から手を下ろしている状態時間が長いことが分かった。

これまでの映像と音声による遠隔コミュニケーション支援では、食事を対象とする研究はなされてこなかった。従来のような共同作業の支援だけでなく、近年では日常生活のコミュニケーションの支援が重要視されてきている。本研究では日常生活の中の食事に着目し、実験によって、共食時のコミュニケーションを対象とした場合でも対面と遠隔でコミュニケーションに違いがあることが分かった。

## 謝辞

最後に本研究を進めるにあたって、指導教員の井上智雄先生に心より感謝申し上げます。  
また、私を支えてくださった同研究室の皆様、特に同じ研究班として活動した方々に深く感謝の意を述べさせていただきます。

そして対面と遠隔食事場面撮影に協力してくださった方々に深く感謝申し上げます。  
ありがとうございました。

## 参考文献

1. 外山紀子：食事概念の獲得：“小学生から大学生に対する質問し調査による検討”、日本家政学会誌、Vol. 41、No. 8、pp. 701-714，1990。
2. 古川大智, 井上智雄, “遠隔 2 者間共食コミュニケーションにおける食事映像の有無の効果,” 電子情報通信学会技術研究報告, Vol. 111, No. 464、pp. 67-72, 2012
3. Cagle, W. B., Stokes, R. R., Wright, B. A., The Picturephone System - 2C Video Telephone Station Set, Bell System Technical Journal, Vol. 50, No. 2, pp. 271-312, 1971.
4. Watabe, K., et al., Distributed Multiparty Desktop Conference System: MERMAID, Proc. CSCW90, pp. 27-38, 1990.
5. Fahlen, L. E., Brown, C. G., Stahl, O., Carlsson, C., A space based model for user interaction in shared synthetic environments, Proc. INTERCHI' 93, pp. 43-48, 1993.
6. Watabe, K., et al., Distributed Multiparty Desktop Conference System: MERMAID, Proc. CSCW90, pp. 27-38, 1990.
7. 石井亮, 中野有紀子, “ユーザの注視行動に基づく会話参加態度の推定—会話エージェントにおける適応的会話制御に向けて,” 情報処理学会論文誌, Vol. 49, No. 12, pp. 3835-3846, 2008.
8. 大塚和弘, 竹前嘉修, 大和淳司, 村瀬洋, “複数人物の対面会話を対象としたマルコフ切替えモデルに基づく会話構造の確率的推論,” 情報処理学会論文誌, Vol. 47, No. 7, pp. 2317-2334, 2006.
9. 古川大智, 井上智雄, “食事の見え方が異なる 2 つの遠隔共食場面と対面共食場面におけるコミュニケーションの違い,” 情報処理学会論文誌, Vol. 54, No. 1, pp. 266-274, 2013.
10. 武川直樹、峰添実千代、徳永弘子、湯浅将英、瀬下卓弥、立山和美、笠松千夏、“3 人のテーブルトークにおける視線、食事動作、発話交替の分析—会話と食事動作はどう制御されるか?—,” 信学技報. HCS, ヒューマンコミュニケーション基礎、Vol. 108、No. 187、pp. 31-36、2008.
11. 武川直樹、峰添実千代、徳永弘子、寺井仁、湯浅将英、立山和美、笠松千夏、“3 人のテーブルトークの視線、食事動作、発話交替から見えるコミュニケーション～銘々皿と大皿料理における行動の比較分析～,” 信学技報. HCS, ヒューマンコミュニケーション基礎、Vol. 109、No. 224、pp. 17-22、2009.
12. 武川直樹、徳永弘子、湯浅将英、津田優生、立山和美、笠松千夏、“食事動作に埋め込まれた発話行動の分析：3 人の共食会話のインタラクションの動作記述,” 電子情報通信学会論文誌. A, 基礎・境界、Vol. J94-A、No. 7, pp. 500-508, 2011.
13. 徳永弘子、武川直樹、木村敦、寺井仁、“3 人の共食会話における発話行為の構造分析、視線方向と発話行為タグによる共食効果の評価,” 信学技報. HCS, ヒューマンコミュニケーション基礎、Vol. 110、No. 459、pp. 55-60, 2011.
14. 中山彰、細田真道、小林稔、“飲みコミュニケーションとパーティコーパスの収集と分析,” 信学技報. MVE, マルチメディア・仮想環境基礎、Vol. 105, No. 256、pp. 83-88、2005.
15. 中山彰、細田真道、小林稔、“飲みコミュニケーションとパーティコーパスの収集と分析,” 信学技報. MVE, マルチメディア・仮想環境基礎、Vol. 105, No. 256、pp. 83-88、2005.

16. 斎藤洋典・喜多壮太郎. “ジェスチャー・行為・意味”, 287. 共立出版, 2002
17. Condon, W. S. & Sander, L. W. (1974). Neonate Condon, W. S. & Sander, L. W. “Neonate Interactional participation and language acquisition” *Science*, 183 (4120), 99-101, 1974
18. 菅原和孝ひとつの声で語ること一身体とことばの「同時性」をめぐって. 菅原和孝・野村雅一, “コミュニケーションとしての身体”, 246-287. 大修館書店, 1996
19. 中村美知夫 “同時に『する』毛づくろい—チンパンジーの相互行為からみる社会と文化 “ . 西田正規・北村光二・山極寿一(編), “人間性の起源と進化”, 264-292. 昭和堂, 2003
20. Kendon, A. (1967). Some functions of gaze direction in social interaction. *Acta Psychologica*, 26 (1), 22-63.
21. Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press.
22. 渋谷昌三: 人と人との快適距離、NHK Books, 1990。
23. [icn.ncu.edu.tw/index.aspx](http://icn.ncu.edu.tw/index.aspx)
24. <http://tla.mpi.nl/tools/tla-tools/elan/>
25. Koiso, H., Horiuchi, Y., Tsutiyu, S., Ichikawa, A., Den, Y.: An Analysis of Turn-Taking and Backchannels Based on Prosodic and Syntactic Features in Japanese Map Task Dialogs, *Language and Speech*, Vol.41, pp.195-321, 1998.
26. 水上悦雄、矢野博之 “対話における間の構造 “ , 人工知能学会研究会資料、SIG-SLUD-A302-08、 pp.43-48, 2003.
27. 米村俊一、宮本勝、中谷桃子。“遠隔サポートにおける映像チャネル導入の効果”。報処理学会研究報告. マルチメディア通信と分散処理研究会報告 2005(3), 99-104, 2005.
28. 玉木秀和、東野豪、小林稔、井原雅行、岡田謙一, “遠隔会議における発話衝突低減手法、” 情報処理学会論文誌、Vol53, No. 7, pp. 1797-1806, 2012
29. 小磯花絵、伝康晴。“円滑な話者交替はいかにして成立するか—会話コーパスの分析にもとづく考察—”。認知科学。Vol.7, No.1, pp.93-106、2000.
30. 榎本美香. “会話の聞き手はいつ話し始めるか : 日本語の話者交替規制は過ぎ去った完結点に遡及して適用される” 認知科学 [1341-7924] Vol.10, No.2, pp 291 -303
31. 大石周平、尾田政臣, 話者間の精神テンポの差がコミュニケーションの円滑化に及ぼす影響: 交替潜時を指標として, 電子情報通信学会技術研究報告. PRMU, パターン認識・メディア理解 Vol.105(534), pp,31-36, 2006.
32. 小宮千鶴子, “相づち使用の実態—出現傾向とその周辺、” 語学教育研究論叢, Vol.3, 1986
33. 楊晶, “中国語と日本語の電話会話における相づち使用の一比較 : 形式と頻度の観点から、” 言語文化と日本語教育 Vol.17 p.1 -13, 1999
34. 大坊郁夫: しぐさのコミュニケーション: 人は親しみをどう伝え合うか、サイエンス社(1994)。
35. 武川直樹、徳永弘子、湯浅将英、津田優生、立山和美、笠松千夏: 食事動作に埋め込まれた発話行動の分析—3人の共食会話のインタラクションの動作記述—、電子情報学会論文誌、 Vol. J94-A, No. 7, pp.500-508 (2011).
36. 大武美香, “3者間共食コミュニケーションにおける発話と食事行動の関係、” 電子情報通信学会技術研究報告, Vol111, No.393, pp.33-36, 2012

37. McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
38. Sellen, A. J, “SPEECH PATTERNS IN VIDEO-MEDIATED CONVERSATION, ACM Conference on Human Factors in Computing Systems, ” pp. 49-59, 1992.
39. 山根智恵 “日本語の談話におけるフィラー”、くろしお出版、東京、2002
40. 中里 収、田本 真詞、菊池 英明、吉村 隆 “課題遂行対話における対話潤滑語の認定” 人工知能学会誌 Vo l . 14, No. 5, pp. 900 -906, 1999
41. 渡辺 美知子 , 広瀬 啓吉 , 伝 康晴, 峯松 信明 “音声聴取時のフィラーの働き : 「エート」による後続句の複雑さ予測” 日本音響学会誌 Vo l . 62, No. 5, pp. 370 -378, 2006
42. 佐々木 寛紀 , 武川 直樹 , 木村 敦 [他] , 徳永 弘子 “非円滑な発話交替時における沈黙の気まずさとフィラーの関係” 電子情報通信学会技術研究報告 Vo l . 111, No. 60, pp. 97-102, 2011
43. 賈キ “小集団討論場面における話者交替の日中対照研究” 世界の日本語教育. 日本語教育論集 Vo l . 18, pp. 73-94, 2008
44. O’ Malley, C., Langton, S., Anderson, A. Sneddon, G.D., Bruce, B.: Comparison of face-to-face and video-mediated interaction, *Interacting with Computers*, Vol. 8, No. 2, pp. 177-192 (1996).
45. Argyle, M.: *Bodily communication*, Routledge, 1988.
46. Moerman, M.: *Talking Culture*, Univ. of Pennsylvania Press, 1998.
47. 榎本美香, 伝 康晴 “3 人会話における参与役割の交替に関わる非言語行動の分析” 人工知能学会研究会資料 言語・音声理解と対話処理 Vo l . 38, pp. 25-30, 2003
48. Furuyama, N.. Prolegomena of a theory of between-person coordination of speech and gesture. *International Journal of Human- Computer Studies*, Vo l . 57 No. 4, 345 -372. 2002
49. 細馬宏通. 話者交替を越えるジェスチャーの時間構造—隣接ペアの場合—. 認知科学, Vo l . 16 , No. 1, pp: 91-102. 2009